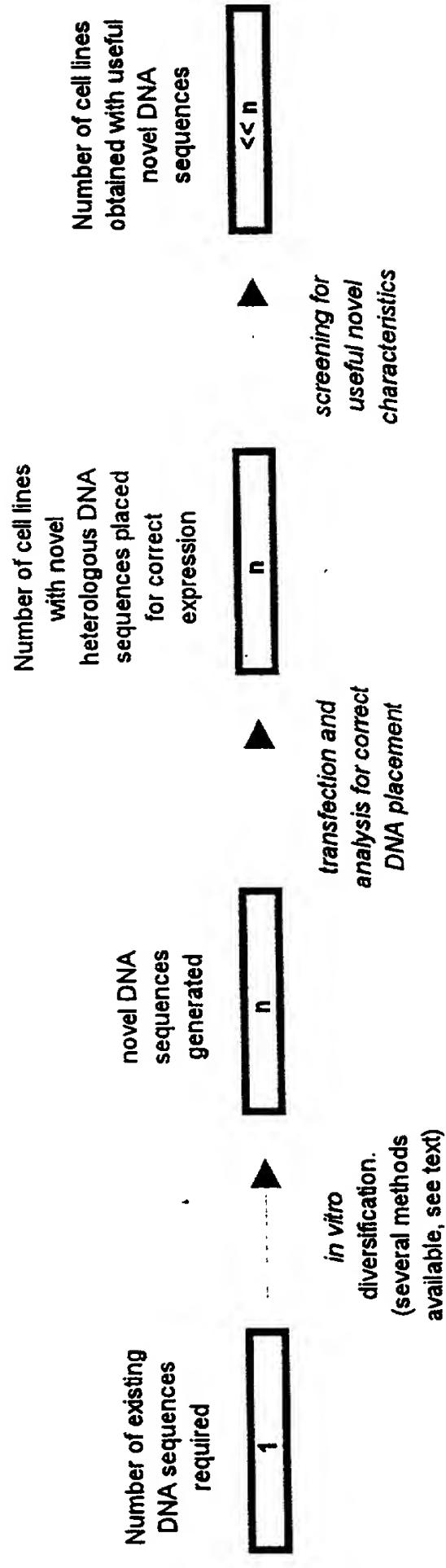


Figure 1 Methods for the diversification of DNA sequences and testing for superior variants

1A existing protocols: Number of transfections needed to generate 1024 new variants: 1024



1B a protocol enabled by the present invention: Number of transfections needed to generate 1024 new variants: 2

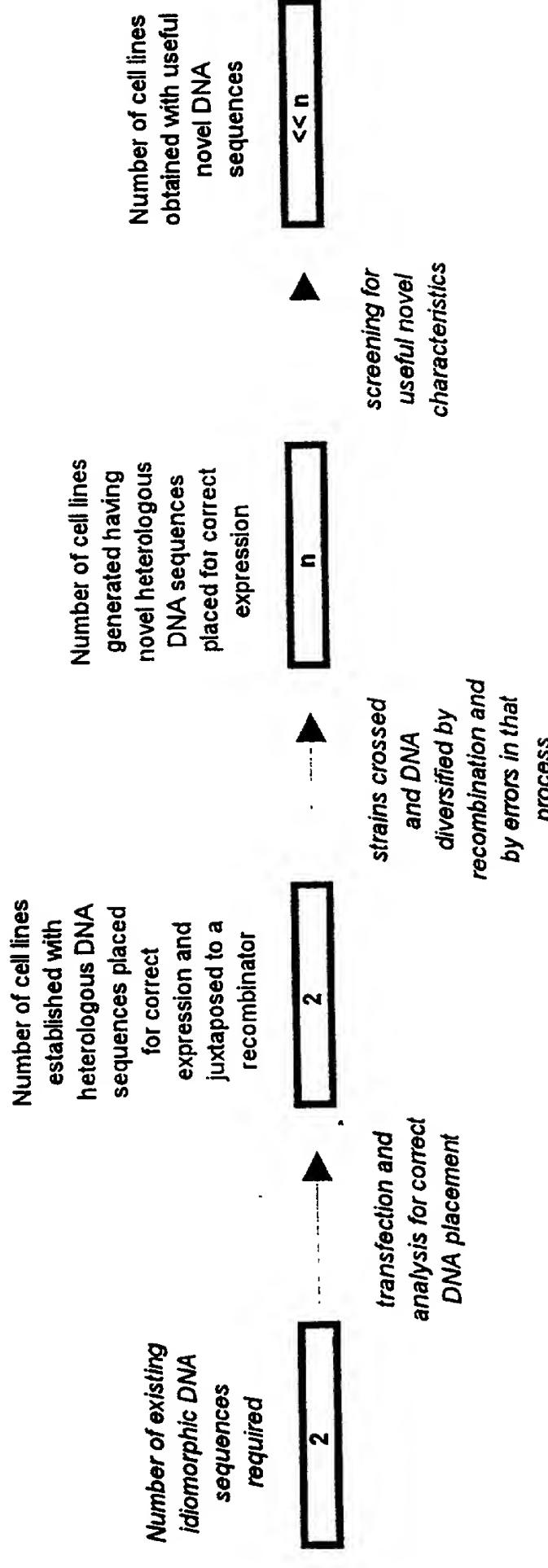


Figure 2 Methods for the diversification of DNA sequences coding subunits of heteropolymeric proteins and testing for superior variants.

The example given for immunoglobulins is for illustrative purposes only and is not intended to limit application of the present invention to this specific heteromeric protein. H = heavy chain genes, L = light chain genes

Figure 2A existing protocol: Number of transfections needed to generate 1024 new combinations: 2048

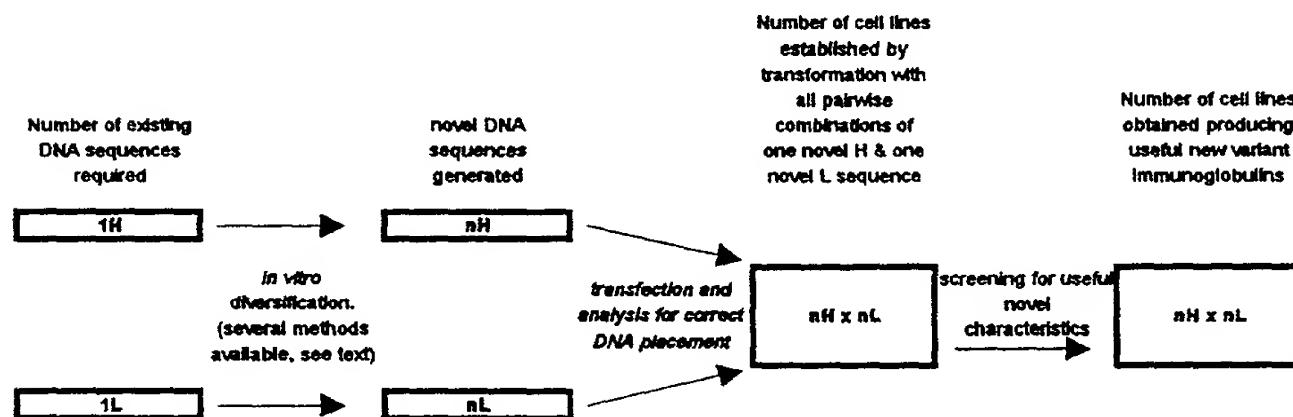


Figure 2B existing protocol using the heterokaryon technology of US Patent Serial No. 5,643,745
Number of transfections needed to generate 1024 new combinations: 64

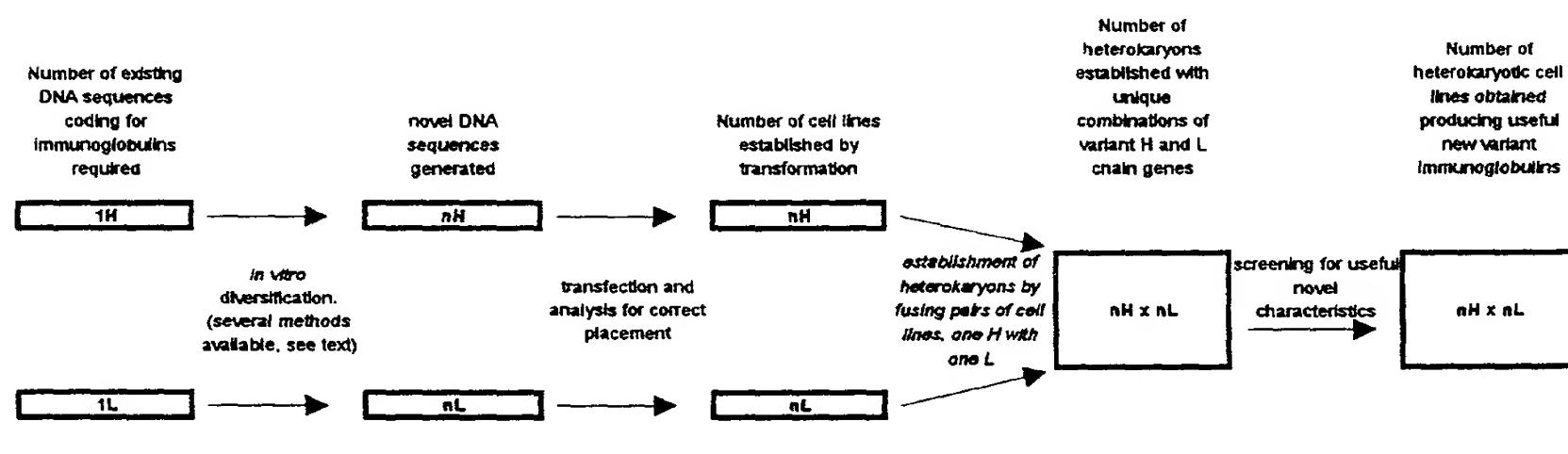


Figure 2C a protocol enabled by the present invention
Number of transfections needed to generate 1024 new combinations: 4

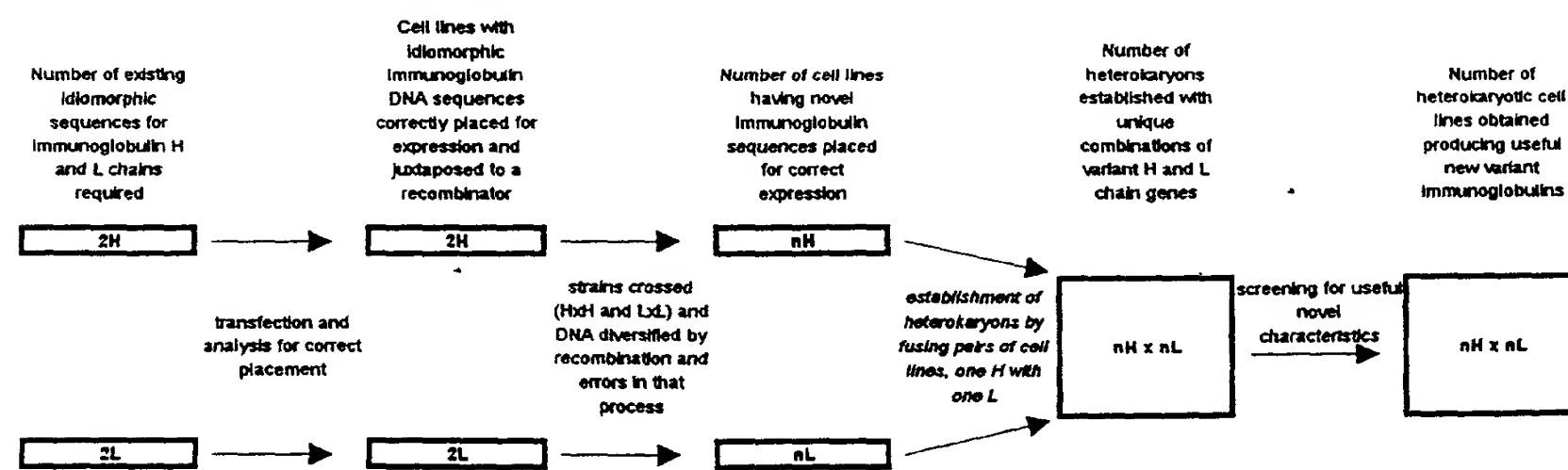


Figure 3 The modified double strand break repair model for meiotic recombination. After H Sun *et al* Cell 64: 1155-1161, 1991

(a) A double strand break (DSB) is made in one DNA duplex. (b) A long 3' overhanging single strand tail is generated either side of the break by resection. (c) One 3' end invades a homologous duplex forming a D loop. (d) the D loop is enlarged by repair synthesis and anneals to the second 3' end. (e) Repair synthesis occurs at the second 3' end and two intermolecular junctions (Holliday junctions) are formed. Resolution of the junctions by cutting inner and outer strands can give rise to non-crossover (f) and crossover (g) chromosomes. If there are base mismatches in the heteroduplex regions (duplex molecules with thick and thin lines) there will be gene conversion. If mismatch repair does not occur there will be post meiotic segregation of new sequence combinations.

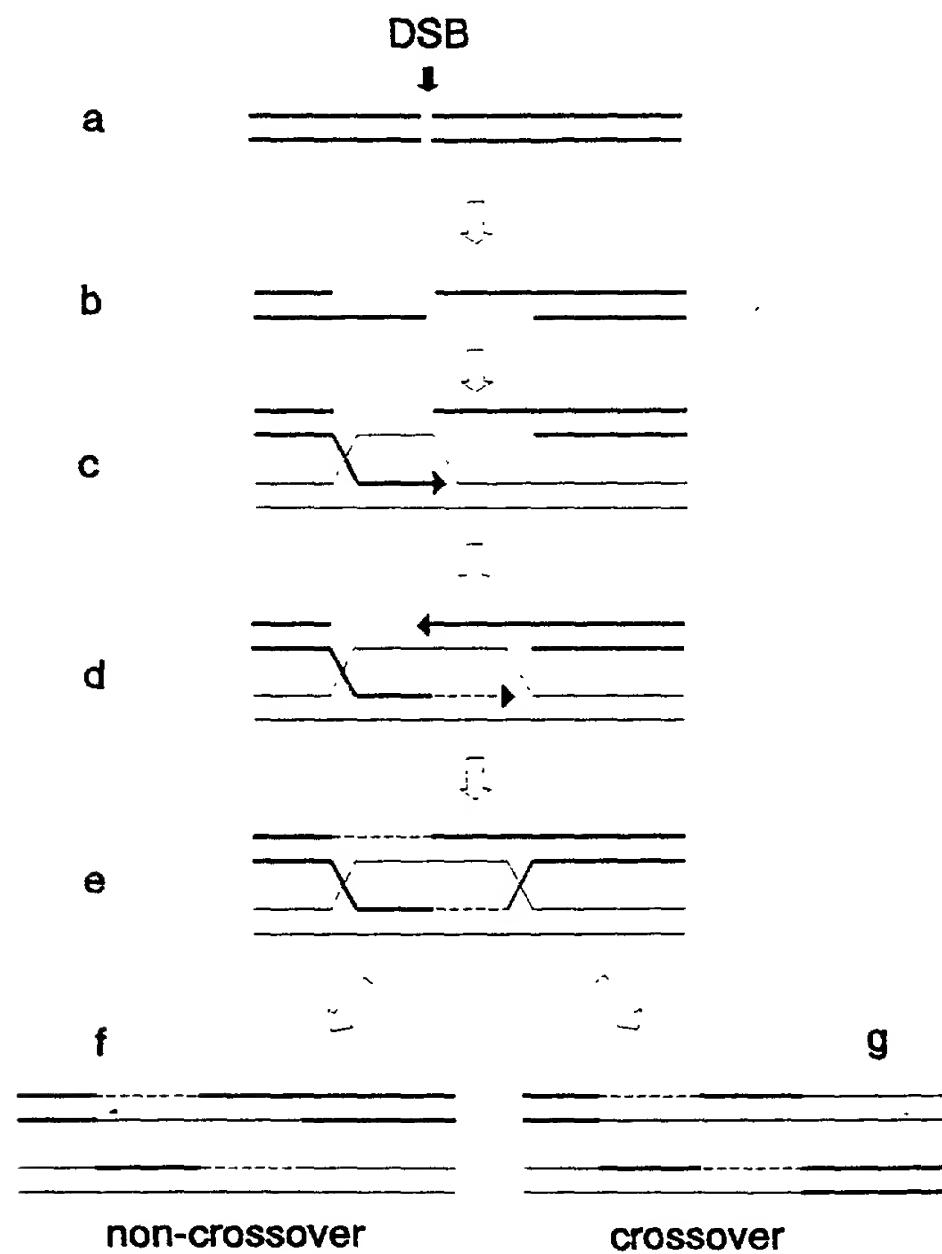


Figure 4 Life cycle of *Neurospora crassa* after JRS Fincham (Genetics, Wright 1983). Microconidia having one nucleus are not shown but can be generated as described in the text. Perithecia and protoperithecia are shown in section.

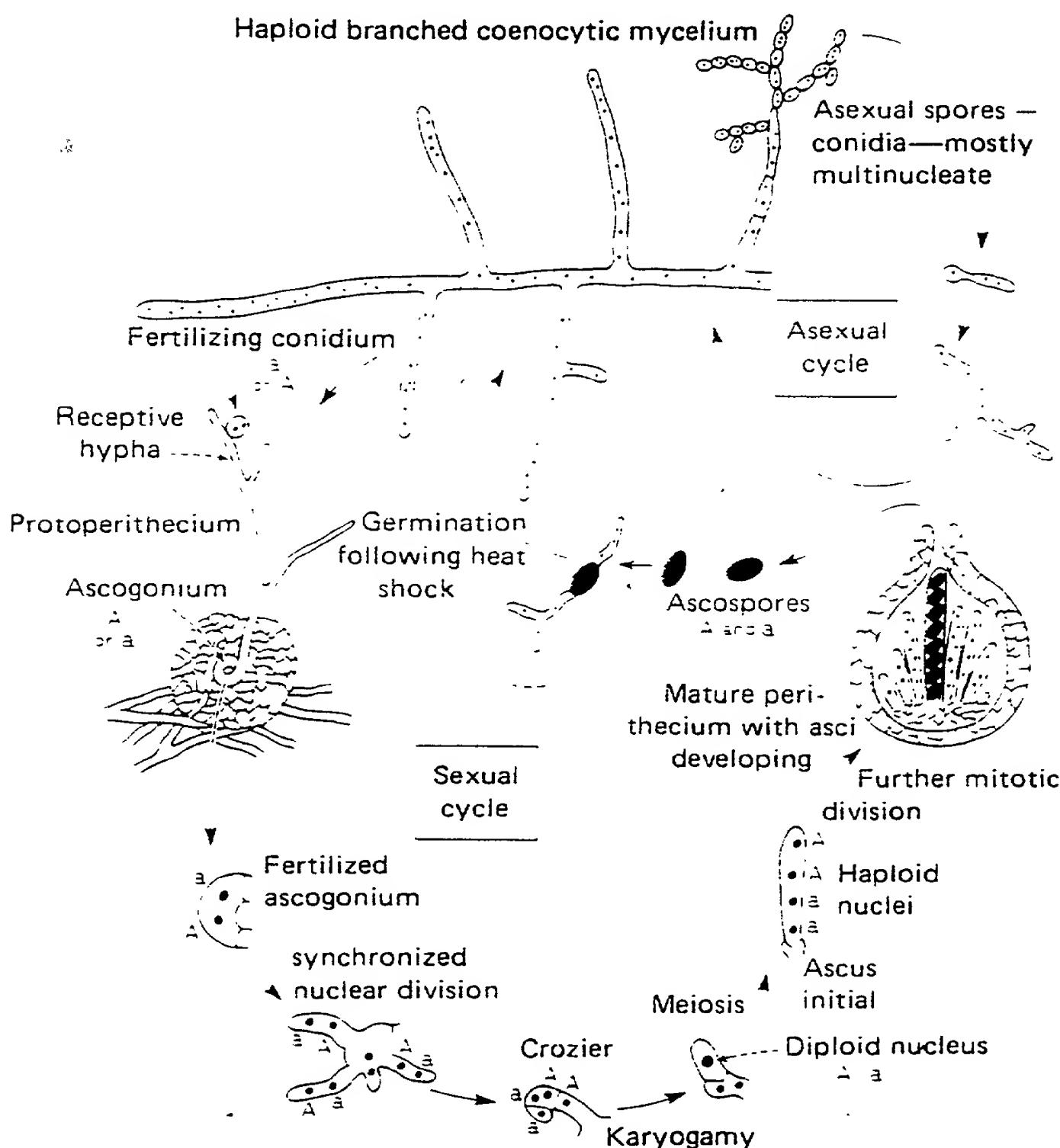


Figure 5 Map of the *his-3*, *cog*, *lpl* region of Linkage Group I of *Neurospora crassa*. Vertical bars, triangles and hairpins show the location of sequence differences that distinguish the St Lawrence and Lindegren wild type strains. The corresponding full DNA sequences are given in figure 7 and figure 8. Vertical slashes indicate one to seven base substitutions per 10 base pairs. Triangles indicate short sequence insertions and the hairpin a 101 base pair inverted repeat transposon fragment present in St Lawrence.

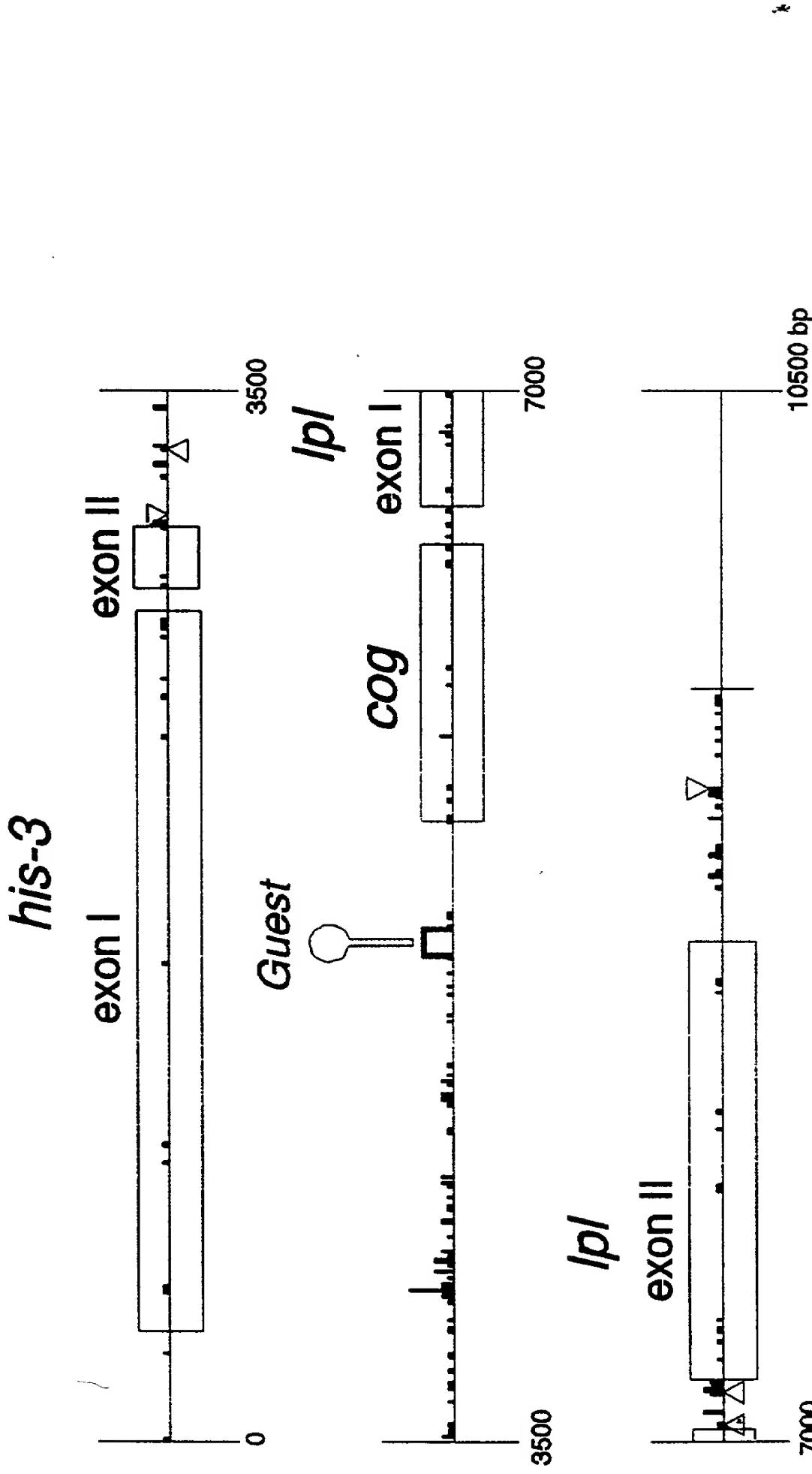


Figure 6. Discontinuity in the parental origin of DNA sequences in progeny from crosses between pairs of *his-3* alleles. In most cases this reflects discontinuity of conversion tracts, in some cases crossovers near the ends of conversion tracts.

Markers are specific DNA sequence differences that distinguish the parents. These were all E (Emerson wild type origin) in the parental strains. Recombinants carry both E and L markers.

Marker position is given in base pairs from the first base of the first codon of the *his-3* gene. Each line of the table shows the parental origin of the markers inherited by one of the progeny.

marker	P	H	P1	K1201	K504	L3	R1	K26	K874	R4	C4	C5	C6	C6/7	C7	C8	C1	C2	C3	C9	D
location	~600	-384	115	179	563	1232	1502	1717	2318	3436	3705	4000	4304	4667	4821	5232	5495	6153	6507		

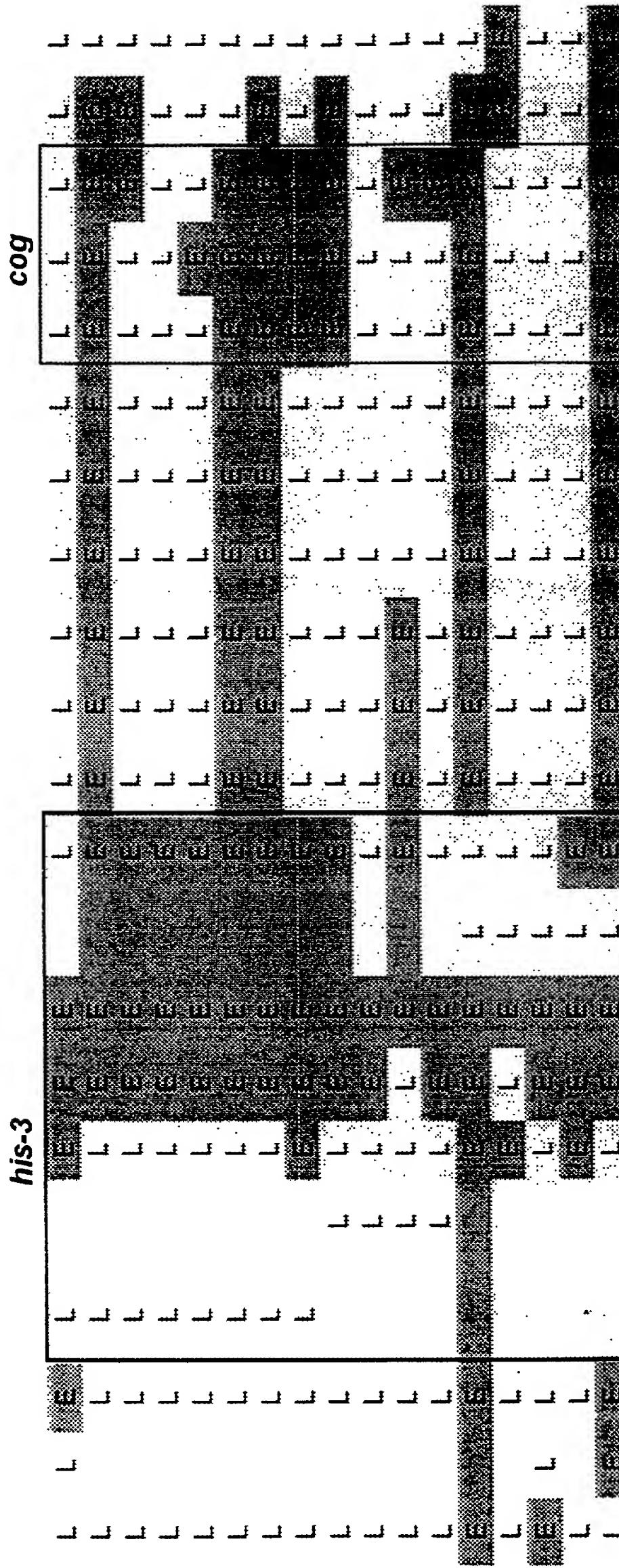


Figure 7 Nucleotide sequence of the *his-3 cog^L lpl* region of linkage group I in the Lindegren wild type strain of *Neurospora crassa*. This differs from that in the StLawrence strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the high frequency recombinator *cog^L* which is active providing the cross in which meiosis occurs is homozygous *rec-2*.

```

1 GATCGCAACT GGAGATCACT CGCACCGTGC CGCAGAACAA GGGCGACGAG CCTCAGGGCA
61 GTTTAGCCTG CCGTAACAGC ACAGACCATA GCTTATTTC ACCTGGGCAG GCAGGGCAGC
121 GCGGCACTGA CATCGGCAAG GCGGCATCAA GCAACCCCTC TGTTGCTTGC CAGCTGCCGG
181 CCAACGTCAG CGGTACAAGG AGAAATCTGG AAGGAAAGAC TTCTGGCACC GACAGGATGG
241 CACGCGGGAA AAAGTTCCAA TGCAATGAGAT GAGGGGCATT TGCATTGCCT CCCGTCACAC
301 TGCCCCGCGAA CCCCAACCCC ACCATAGCGT CTGTCGATAC ATGGAGCGCG AAGTCGAGAA
361 ACCTGTAATT CCTGGTAACT TTCAGGTACA CAGTACGTAC TGATCCTGGT ATCAAACCTT
421 GCCTGCCGAG TTTTCGACGG AAAGAGGTGT GAATTGTGAA AGAGTCATAC CAAATCACCC
481 GATTTTCATA AAGCCCAGT CTTTCTGTA CATAAGCGAC ACTCGAAGCG GGCCTCATCT
541 TCATAGCCTG ATAGCTTGTA ATACTCCATC CTCGTATCTC ACTTGACCTT GAGTTCAACC
601 CCACGTCAGA CTTCACCCGA CACATCGACG GATTGGGGAA CAGCACAATA CCTGAAAAGC
661 GAGAAAACCA AACAGAGGAA AACACCATGG AGACAACACT TCCCCTCCCC TTCCCTCGTCG
721 GTGTCAGTGT TCCTCCCGGA CTGAATGACA TCAAGGAGGG CCTCAGGCC GAGGAAGTCT
781 CGTGTCTTGG CTGCGTCTTC TTGAGGTCA AGCCCAAGAC CCTTGAGAAA ATCGTGCAT
841 TCCTCAAGCG TCACAATGTC GAATTTGAGC CCTACTTCGA TGTAAACAGCC CTCGAGTCTA
901 TCGATGATAT TATCACTTT CTGGACGCCG GCGCCCGCAA GGTGTTGTC AAGACCGAGC
961 AGTTGGCCGA CCTCTCCGCA TATGGCTCCC GCGTTGCCCG CATTGTCACT GGAAGCAGCG
1021 CTGCTTGCT TTCCCTCCGCC ACCGAGAGCG GCCTTTGCT CTCCGGCTTC GATCAGACTG
1081 CCTCCGAGGC TGACAGTTT CTGGAGGAGG CCAGAGACAA GAAAATTACC CCCTTCTTCA
1141 TCAAGCCCGT TCCTGGGGCC GATCTCGAAC AGTTCATCCA GGTGCGCGCC AAGGCTAACG
1201 CCATCCCCAT CCTGCCATCC ACTGGCTTGA CAACAAAGAA GGACGAGGCC GGAAAGCTTG
1261 CCATCTCCAC CATCCTCTCG AGCGTCTGGA AGTCTGACCG TCCCAGATGGT CTGCTCCCCA
1321 CCGTTGTCGT TGATGAGCAC GACACTGCTC TGGGTCTGGT CTACAGCAGT GCCGAGAGTG
1381 TGAACGAGGC CCTCAGGACA CAGACTGGTG TCTATCAGAG CCGGAAGCGC GGTCTCTGGT
1441 ACAAGGGTGC TACTTCCGGA GACACTCAGG AGCTCGTCCG CATCTCGCTT GACTGCGATA
1501 ACGATGCTCT CAAGTTGTC GTGAAGCAGA AGGGTCGTTT CTGCCACCTC GATCAGTCCG
1561 GCTGCTTGG TCAGCTAAA GGCCTTCCCA AGCTCGAGCA GACTTTGATT TCGAGGAAAC
1621 AGTCTGCCCG CGAGGGCTCC TACACTGCCG GTCTCTTCTC CGATGAGAAG CTAGTCCGGG
1681 CCAAGATCAT GGAGGAGGCT GAGGAGCTCT GCACCGCTCA GACCCCCCAG GAAATCGCCT
1741 TTGAGGCTGC CGATCTCTTC TACTTGCTC TTACCAAGGGC CGTTGCTGCC GGCCTTACTC
1801 FTGCGATAT CGAAAGGAGC CTTGACGCCA AGAGCTGGAA GGTCAAGCGC AGGACTGGAG
1861 ATGCTAAGGG TAAGTGGGCT GAGAAGGAGG GCATCAAGCC TGCGGCGTCC GCTCCCGCTG
1921 CCACTTCGGC CCCTGTCAAC AAGGAGGCCG CCCAGGAGAC CACCCCTGAG AAGATCACCA
1981 TGAGACGTTT CGACGCCTCC AAGGTCTCTA CCGAGGAGCT CGATGCTGCT CTCAAGCGTC
2041 CTGCGCAAAA GTCGTCCGAT GCCATCTACA AGATCATTGT CCCCATCATC GAGGACGTCC
2101 GCAAGAACGG CGACAAGGCT GTTCTGTCGT ACACTCACAA GTTCGAGAAG GCTACCTCTC
2161 TTACTAGCCC CGTCCTGAAG GCGCCCTTCC CCAAGGAGCT TATGCAGCTC CCTGAGGAGA
2221 CCATTGCTGC CATCGACGTG TCCTTCGAGA ACATCCGCAA GTTCCACGCC GCCCAGAAGG
2281 AGGAGAACCC CCTCCAGGTC GAGACCATGC CCGGTGTTGT CTGCAGCCGT TTCTCTCGTC
2341 CCATCGAGGC CGTCGGCTGC TACATCCCCG GCGGTACCGC CGTTCTCCCC AGCACTGCC
2401 TTATGCTGGG TGTTCCCGCC ATGGTCGCCG GCTGCAACAA GATTGTGTTT GCCTCTCCTC
2461 CCCGCGCCGA CGGAACCATC ACTCCCGAGA TTGTCCACGT CGCTCACAAG GTTGGGGCCG
2521 AGTCCATCGT GCTTGCCTGC GGTGCCAGG CCGTAGCTGC CATGGCCTAC GGCACCGAGA
2581 GCATCACCAA GGTGACAAG ATTCTCGGCC CCGGTAACCA GTTCGTCACT GCTGCCAAGA
2641 TGTTGTCAG CAACGACACC AACGCTGCCG TTGGGATTGA CATGCCCGCT GGCCCGTCCG
2701 AGGTGCTGGT CATCGCTGAC AAGGACGCCA ACCCCCGCTT CGTTGCCTCG GATCTCTGT
2761 CCCAGGCTGA GCACGGCGTT GACAGTCAGG TCATCCTGAT CGCTATTAAC CTCGACGAGG
2821 AGCATCTTCA GGCTATTGAG GACGAGGGTC ACCGTCAAGGC TATGGAGCTT CCTCGCGTCC
2881 AGATTGTCGG TGGCTCCATC GCCCACTCGA TCACCGTGCA GGTCAAGACC GTCGAGGAGG
2941 CCATGGAGCT CAGCAACAAG TACGCTCCTG AGCACTTGAT CCTCCAGATC AAGGAGGCCG
3001 AGAAAGCTGT CGATCTTGTC ATGAACGCTG GTAGTGTCTT CATTGGCGCT TGGACTCCTG
3061 AGTCCGTTGG CGATTACTCT GCTGGTGTAA ACCACTCGCT GCGTAAGTTA CATATCATAA
3121 ATAGCCCCGC TTCACAGATT CTTCTGCTAA CGTCAAGACA CATAGCTACC TATGGTTTTG
3181 GCAAGCAGTA CTCTGGCGTC AATCTCGCCT CGTTCGTCAA GCACATTACC AGCTCCAAT
3241 TGACTGCCGA GGGTCTCAA AACGTGGCC AGGCTGTCA GCAGTTGGCT AAGGTTGAGG
3301 AGCTCGAGGC TCACAGAAGG GCGGTCAAGCA TCCGTCTTGA GCACATGAGC AAGAGCAACT

```

Figure 7 continued

3361 AGACGGAAAT TCTTTTCGA AGTTGCAAAA AAAACAAGAA CAAAAGGATG TAGTGGGTTG
3421 ATGTATATCT GGGTCATTG GGGCACATAG AGTAATGATA ACGAGTTTG GACATTGTAC
3481 TGTTCTGTAC AGGCTGAAGA TCAGTACATG AATCTGTTGG TAAGTGTAGA GACCCAAACG
3541 TCCCTTGAGT TTTCTCCCT GTTCCAGAGA GGTGCTCGTC CCTGGGTGTT TATTTTCATT
3601 ATTACATCAA CCTTTTATT TATTTTATT TTTTATTAC TTTTTTTCC TTTTTTCAG
3661 ATCATGCGTA CATGAACGGG GGAAGCACAG ACGATCGAAA CGTGGATGTC ACAATGTCGC
3721 TGCAGTGATG CTGCATTGCA TGAAGCGCCC ATCTCAATAT ACTTGCAGTC TTGCGCGTTG
3781 CACGTGAAC TCCAAACAA CGAATAAAA GACGGCGAAA AATGAAGATA AAAAAAAACC
3841 ATAATAAAA TCGGAGGGAG TGTGGAAAT GTTTCTTT AGCATTAGA CCCCATAGCC
3901 GTGCACGCC GGGTACAGAC AGGTTCATCG ATGTTGACAT TGACTGGAC ACCAGGTCTA
3961 TCTATTTCAT CTCCCTGTCCT CTACCATACA TCGGGACATC GGACATCTCG CTGTACCCCC
4021 CACACCCACA AAGTCTTATA AAAGCGCCAC ACCCGAGGAG GTTCGGTCGG CCCCACGAAC
4081 TCCGTGCCTC CCTGCCTGTT TACAGGGACC GAACGCTGGA GAAGCTTAGT TTCCGTACAT
4141 CCGGCCTACC CGAGCAGGAA AAGGGACAGC TCATAGGCAGA GGAGGGATTG GAAGATGGGG
4201 ACATTTGGA TGATTGAGA GGAGGAACTA GGTACTGTAT CATGATAGTT CGGGGCAGCA
4261 TCTTGGCTGG GACATTGTA ATACCTCGAT ATGATGAAGT GGGAGGGAGT TTTTCATGT
4321 CTTGCCAAG TCCCACTAAT CTTTTTTTT TTTTGTACCA ACACCCAAGA TTCTGGAGAAT
4381 AGTGTAAAGGA TTCGCATTCA CAAGTGGAAAG TCTGAGGATC TTTTATATC TTTGTCTTCC
4441 GCGGACTGTT AACGATCTA CAGCGAGCGA GCGAGCGTC GGATGCGCTG ATCTGATAGG
4501 TGCAATATAC GGCGCTTTC TCCGGTCGTG TAGTGTAAAGC TCTGCGGCA TAGTAGTACA
4561 CTAAAAAAAC CCTTGCAATT CATGATCTGC TTGCTATTCA TTCCGAGTTA TTTCAGTGGT
4621 CACATTCGA GATTACACAGC CATCCATCCA TATGGAAAAA TCCATTCCCA TGCTTCCCTCC
4681 CCCCCACTAT GTATGTGACC ACACGCTGCT GTCAGAATGC CAACGGCTC AGGTACCCCTC
4741 GTCCGACTGT TTGGCATGGA GTTACATACA CTACTAGTGT AGCCCCGGGC CAAGCTACCC
4801 CGTCAAATCT ATACATATCT ATAATGGGTT TCAGGTGTT CGTTCGCTG CAATCAAGTT
4861 TGAAACATCA CTGGGGCCGT TGGACGGTGT ATTAGACCAT TGGCTCCCTC AGCTGGCGGC
4921 TGGGCGGTTG GGTGGCAAT AACGGGACTG GACTTGAGAG GGACGAGGAG AGTCGGTTGG
4981 CTGCCTACAC TACACTACAA GCGTTCCCAC CTAACCGACG AGTCCCCTT TCCATTGTG
5041 TGCCCTAACC ATCATCTAGG GATGTCAAGG TTTGGCCCGA TCAGGGTATG TTTGGTTGAC
5101 TGGTGTATG TCTGATTGGG TACATATCAT GGTAGGTGTC TCGAGAACAG TAGAGTACTC
5161 GGGCTAGCG TTTGGATGAT TACGCGAGAT ATGAGTTGTA GGCCGCCATG CAGTTGCTTG
5221 CCCATAAGCA GAAGTTGCTT TGGGATATAT TTCTCGTCTT TCAAAGGTCA CGAGGTCTG
5281 GGACGAGCGG CATGCCATC CAAAGGGTTG AACATGAGAA ACCGGAATGG CCTTGCGTT
5341 GAAATACAAA AAGTCAAGAA TAAAATCGCT TGAGGATAGG GACGTGGAAG CAAGCAAATA
5401 TGGTAAGGGA GGTACTGCTA TGTAGGTGCT CAGCAAACGT CCAATTCTT GGCCCCAAG
5461 CAGCAGTTG CTGTCAGTGC TGCTCGTGC AGCCTGGTA GTGGAACCTA AACTGCTAAC
5521 ACAGCGCAAG TGCGCATGTA AAGATATTGT GGGAGGATCT GTATGGATGG ATGAGATTAC
5581 TGCTTGGTGT TGGTTGCGAG GCACTGCGGC TGTAGGCTT TGCTGTGCC CGTTGACGA
5641 AGAAAATACGC GGAACATATAA ATTGGATAACC TAGACTTAAT GCCTATGGGA GGTATCTACC
5701 GACGTAGCCG ACGGATTCTA GCAACATCCC GACTTGCTT GTAGTGTACT ATGATAGCAG
5761 CACAGTGGGG TGTGCTCCT TGTGAGGATG GGCTCTTTT TTTTTTTCC CCCTTCCCTA
5821 GGGCGTTGAC TGGACTTGCT CTATCGTCC CAAGGTAGGT GCCCGTCATC GATTTCCCA
5881 AGCCGTCTCC CGCCAGATTG TCGTCATAGT GTCATGATGA CCTCGTCGC TGGGCGTGC
5941 TGGTTACGGG GAGCTGGGAC CGCTAGGCCT CAGTGGTTGT GCCATTTCAGC GTGGGTGTGT
6001 GGAGTAGCGG TAGAGGGCGCT TGGAAAGTTGT GCTAGGGAA ACCCTGGAAT ATCTTGTACC
6061 CTTCGATTCC TTCTCGGGCT GCCCATGTGC TGAGGTGATG CCGGGATCT GGCGCCAATC
6121 ATCCATTGAG GTTCCCGCAG CTTCCCGGT CGCGCGCGG GCGCAGTTGC TCACAGGACA
6181 CACCTAGACG CAGGGGCACA GGGGCACCGT TTGGTGTGCA ACTGGGTACC TGGTAGCTGT
6241 AGCAAGCACT CCACCGTCTG TGCAATCCCC CAATCCACGG CAGGAACCTA GCACCGCCGC
6301 GGCACCGAGT GAGCGAATCC ATCCGCATTG GATCCCAATT CTTGCCCTTG CCATCCTCT
6361 TTCTTCCAC TTGGCGCAAC CAACACCTCC CTTGGTCTGG GTACTCGTGT TGATCTCAC
6421 TCTCTTTT TCTTGGCGA CCGACTTTT ATATCCGTCC TTGCTTCCCG CTGGCGTTG
6481 TCGTTCTTC TACAACCTAC TTCCGTTCAT TATCCCTTT CTTGGTTCGG TCGAGGACCC
6541 AAAAACAGAA CAATTCCGGC TCTTCCAGGT GGCTTGGGTG CGACTGTTTA GCTCTTGACC
6601 ACTAGCCGCT TACCTTCTCT TGATGTTTAT ATTTGGATAT CATTGAACTA CTCTTCTTG
6661 AAACGGCAGA CGAACGGAAC AGTCCCTACG GTTTATTAGC GATATACGTT GTACTGATAT
6721 CCTGAGCAAG AAGAGGCAA TTATCAATT TGCACTCTCC ATCGTCGCTG CTCATCGCAG
6781 CTCCCTTGCT CGCCAATGTA TCGGCCGAAC CGATTAGGAT ACCCCAAACGC GATGTTCTCC
6841 GTGGTATCAA CATCACAGCA ACTTGCCTT CGAGCACTAC CGAATTGCGC CAGCGGTGGA
6901 TATGCCCTG CCGTTGTAGA CTGTCCTAAC ACCAAGCCGA CGCTCCGGAA GGCGCGTGGAT
6961 TTGTCGAACG AGGAGAAGAA CTGGTTGTCG ATCCGGAGGA AGAACACCAT CCAGCCCCATG
7021 AGGGACCTAC TGAAGAGGGC CAACATCACT GGGTTCGATT CCGAAACTTT CATGAATGAG
7081 GCCGCCAAC ACGTCTCGCA ACTGCCAAT GTCGCCATTG CCATTTCAGG AGGCAGGCTAT
7141 CGTGCCTCA TGAACGGCGC CGGCTTCGTT GCTGCTGCGG ATAACCGGAT TCAAATACCC
7201 ACGGGCGCAG GTGGTATTGG AGGCTTGTG CAGTCCAGCA CATATTGTA TGTAAAACCA
7261 TGCCTTCTTG TGGTTCTTCT TATCTCGTT TCGAGTGTCA ACTGCGCCAG TTCGACGTTG

Figure 7 continued

7321 GGC GGCTGTG GACGACCTTG CTGGTGAACA TGTCTGGAC TCCATGCCCC TTTTTTCGTT
7381 CCCTAAAATC CCAAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAATTGAG
7441 GACC GTGACT GTAAATTGCT AACGCAACTC TAGGGCCGGA CTTTCTGGTG GTGGCTGGCT
7501 TGTCGGCAGT TTGTTCTCCA ACAACTTCAG TAGCATTGAG ACCCTGCTGA GCGAGAACAA
7561 AGTCTGGAC TTTGAGAAGT CCATCTTAA AGGACCCAAG GAGGCTGGCC TTAGTACTGT
7621 CAACCGTATC CAGTA CTTGGT CCGAAGTGGC AAAGGAAGTT GCGAAGAAGA AGGATGCTGG
7681 CTTCGAGACA AGTATAACAG ACTACTGGGG CCGAGCATTG AGTTACCAAC TGATCGGAGC
7741 CGATATGGGC GGCCC GGCTT ACACCTTCTC CAGCATTGCC CAGACCGACA ACTTCCAGAA
7801 GGCGAAGACG CCGTCCCTA TTCTGGTAGC TGACGGCCGC GCGCCTGGAG ACACCATCAT
7861 CTC CCTCAAT GCTACCAACT ACGAGTTCAA CCCGTTGAG ACGGGTAGCT GGGACCCGAC
7921 CGTCTATGGC TTTGCGCCGA CCAAGTACCT CGGCGCCAAC TTCAGCAACG GCGTGTCCC
7981 ATCGGGAGGC AAGTGC GTTGA AGGGTCTCGA CCAAGCCGGC TTCGT CATGG GCACCAAGCAG
8041 CACGCTCTTC AACCAAGTCC TTTTGGCCAA CATCTCCAGC TACGACGGTG TTGCCAGACG
8101 TGCTCATCGA GGCGT GACT TCTGTCCTCA AGGAAATCGG CGCCAAGAGG ACGACGTCTC
8161 CCAAATCATC CCTAATCCGT TCCTGGACTG GAACAACCAG ACCAACCCCA ACGCCGACAC
8221 GCTCGAGCTC GACCTGGTCG ACGGCGCGA AGATCTGCAG AATATTCCGC TCAACCCGCT
8281 CACCCAAACCC GTGCGCGCCG TCGACGT CAT CTTCGCTGTC GACTCGTCCG CCGACGTGAC
8341 AAAC TGGCCC AATGGCACCG CCCTGCGCGC CACCTACGAG CGCACTTTG GCTCTATTTC
8401 CAACGGGACA CTCTTCCCCCT CGATCCCCGA CGACTGGACG TTTATAAAC TAGGCCTCAA
8461 CAACCGCCCCC TCTTCTTCG GCTGCATGT TAAGAACCTT ACCTTGAACG CCAACCAAA
8521 GGTTCCCCCCC TTAATCGTCT ATGTCCCCAA CGCGCCCTAT ACCGCGCTGA GCAACGTGTC
8581 CACCTTCGAT CCGTCATACA CGATGTCTCA GCGCAACGAC ATCATCGGCA ACGGATGGAA
8641 CTCAGCCACG CAGGGAAACG GCACGCTGGA TTCGGAGTGG CCCACTTGC GCGCCTGCG
8701 GGTTATCAGC AGGAGCTTAG ATCGGTTGGG CAGGCAGACG CCAGCCCGT GCAAGACTTG
8761 CTTTGACAGG TATTGCTGGA ATGGCACAGT GAACTCCAAA GATA CGGGGGG TTTACATGCC
8821 TGAGTTCAAG ATTGCGGATG CGCATGCCCT GGACTCGGGT GCTGTTGCTA TCGGAAAGAT
8881 GGTGAATGTC TGGTCGTGCG TTGTGGTGGG AGTTGTGGCG GCTACTTTGT TGTTGTAGGG
8941 GTAGGGGAGA CGTGATGATA TTCCAGTCTG ATGAAGTTGA GACTGGACTG GAGATGCCA
9001 AGGATGCGGA GGGAAAGGAA TCGGTGGTGT TAATGTCATG ATGGATGAAG AGTCATGGAT
9061 CATGGAACGA CGGGGCGGGG ATATTGGATG ATGGATATAC CACACTGCAT GCATGCTCTA
9121 TTGATAGTAT GCTTGGCAT TTACGTTAA CAATCAATTG CTCCATCCTG ATGTTCTATC
9181 TTTTCGACA ATGGATTGAT ACTACTCCTG TTGCTTCGCT CTTGAGGTTG GAAGGACTTG
9241 AGGTTGGAAG GACTTGAGGT TGTTGTTCT GAGGGAGGTT ATCGAAGTAT CATCTGTGCT
9301 GATGCCGATT GATAGACTGT CCTCTTCTTC GAGGCAACGA ACGGTGGAT GAGCCTCTT
9361 AATCATGATG CTCAGTGCCA CAAAAAGGCT CCAGCACAGC TGCCCACACC TTTCTTGCT
9421 CGCCGTTCCCT TCCTTTTCT TTTCCCCTGT TTCTTTCTT CTTTCCATC TCATCCGTA
9481 CCAGAGTGCC CACCGGGTAT ATATATTACC TCCTTGCGCG TTCTCCTTTG ACCAATAAAT
9541 CGCTTGGTCG AGTGGCGTAA CGGTTTACCG TCTACACTTA TCACTCAAAC CAAACCAAAC
9601 CATCGAAGAA GTGACCTATC GGTTCGAGGG AACGGTGATG TTCTTACGAC CAAGTTAAC
9661 CAAAGAGCGT TCCACATCGT TGAACCGTCT CCTCCAGTTG GATCTGTTA ACTTCCGCAG
9721 CGACTGAAGA AGGTATCACT TTTTTTTGG TTCCAAAAAA AAAAAAAA ATTAC

Figure 8 Nucleotide sequence of the *his-3 cog^E lpl* region of linkage group I in the StLawrence wild type strain of *Neurospora crassa*. This differs from that in the Lindegren strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the weak recombinator *cog^E* and also the remnant of a transposable element *Guest* within the replaceable sequence 3' of *his-3*. StLawrence strains carry *rec-2⁺* which prevents the initiation of recombination at *cog*.

```

1 ACCGGGAATC GTAGCGGGCG CTAAGGCCAA GCCGCGGCAC GGGTCACTGA CCCAATGCAG
61 CGCATTCCGT CAGCACTGA AGTGGATGTA CAAGTACATA GTAGTAGATC GCAACTGGAG
121 ATCACTCGCA CCGTGCCGCA GAACAAGGGC GACGAGCCTC AGGGCAGTTT AGCCTGCCGT
181 AACAGCACAG ACCATAGCTT ATTTTCACCT GGGCGGGCGG GCGACGGCGG CACTGACATC
241 GGCAAGGCAG CATCAAGCAA CCCCTCTGTT GCTTGCCAGC TGCCGGCCAA CGTCAGCGGT
301 ACAAGGAGAA ATCTGGAAGG AAAGACTTCT GGCACCGACA GGATGGCACG CGGGAAAAGT
361 TCCAATGCA TGAGATGAGG GGCATTTGCA TTGCCTCCCG TCACCCAGTG CGAACCCCCAA
421 CCCCACCATA GCGTCTGTCG ATACATGGAG CGCGAAGTCG AGAAACCTGT AATTCTGGT
481 AACTTTCAAGG TACACAGTAC GTACTGATCC TGGTATCAAA CCTTGCCCTGC CGAGTTTCG
541 ACGGAAAGAG GTGTGAATTG TGAAAGAGTC ATACCAAATC ACCCGATTTC CATAAAAGCCC
601 GAGTCTTTTC TGTAACATAAG CGACACTCGA AGCGGGCCTC ATCTTCATAG CCTGATAGCT
661 TGTAACTAC TACCTCGTA TCTCACTTGA CCTTGAGTTC AACCCCACGT CAAACTTCAC
721 CCGACACATC GACGGATTGG GGAACAGCAC AATACCTGAA AAGCGAGAAA ACCAAACAGA
781 GGAAAACACC ATGGAGACAA CACTTCCCT CCCCTTCCCTC GTCGGTGTCA GTGTTCTCC
841 CGGACTGAAT GACATCAAGG AGGGCCTCAG CCGGGAGGAA GTCTCGTGT TTGGCTGCGT
901 CTTCTTCGAG GTCAAGCCCC AGACCCTTGA GAAAATCCTG CGATTCTCA AGCGTCACAA
961 TGTGAAATTG GAGCCCTACT TCGATGTAAC AGCCCTCGAG TCTATCGATG ATATTATCAC
1021 TCTTCTGGAC GCCGGCGCCC GCAAGGTGTT TGTCAAGACC GAGCAGTTGG CCGACCTCTC
1081 CGCATATGGC TCCCGCGTTG CCCCCATTGT CACTGGAAGC AGCGCTGCTT TGCTTCTCTC
1141 CGCCACCGAG AGCGGCCTT TGCTCTCCGG CTTCGATCAG ACTGCCTCCG AGGCTGCACA
1201 GTTTCTGGAG GAGGCCAGAG ACAAGAAAAT TACCCCCCTTC TTCATCAAGC CCGTTCTGG
1261 GGCGGATCTC GAACAGTCA TCCAGGTCGC CGCCAAGGCT AACGCCATCC CCATCCTGCC
1321 ATCCACTGGC TTGACAACAA AGAAGGACGA GGCGGGCAAG CTTGCCATCT CCACCATCCT
1381 CTCGAGCGTC TGGAAAGTCTG ACCGTCCCGA TGGTCTTCTC CCCACCGTTG TCGTTGATGA
1441 GCACGACACT GCTCTGGTC TGGTCTACAG CAGTGCAG GAGTGTGAACG AGGCCCTCAG
1501 GACACAGACT GGTGTCTATC AGAGCCGGAA GCGCGGTCTC TGGTACAAGG GTGCTACTTC
1561 CGGAGACACT CAGGAGCTCG TCCGCATCTC GCTTGACTGC GATAACGATG CTCTCAAGTT
1621 TGTGTTGAAG CAGAAGGGTC GTTTCTGCCA CCTCGATCAG TCCGGCTGCT TTGGTCAGCT
1681 CAAAGGCCTT CCAAGCTCG AGCAGACTTT GATTTCGAGG AAACAGTCTG CCCCCGAGGG
1741 CTCCTACACT GCCCGTCTCT TCTCCGATGA GAAGCTAGTC CGGGCCAAGA TCATGGAGGA
1801 GGCTGAGGAG CTCTGCACCG CTCAGACCCC CCAGGAAATC GCCTTGAGG CTGCCGATCT
1861 CTTCTACTTT GCTCTTACCA GGGCCGTTGC TGCCGGCGTT ACTCTTGCCT ATATCGAAAG
1921 GAGCCTTGAC GCCAAGAGCT GGAAGGTCAA GCGCAGGACT GGAGATGCTA AGGGTAAGTG
1981 GGCTGAGAAG GAGGGCATCA AGCCTGCGGC GTCCGCTCTC GCTGCCACTT CGGCCCTGT
2041 CACCAAGGAG GCCGCCAGG AGACCACCCC TGAGAAGATC ACCATGAGAC GTTTCGACGC
2101 CTCCAAGGTC TCTACCGAGG AGCTCGATGC TGCTCTCAAG CGTCCTGCGC AAAAGTCGTC
2161 CGATGCCATC TACAAGATCA TTGTCCCCAT CATCGAGGAC GTCCGCAAGA ACGGCGACAA
2221 GGCTGTTCTG TCGTACACTC ACAAGTTCGA GAAGGCTACC TCTCTTACTA GCCCGCTCCT
2281 GAAGGCGCCC TTCCCCAAGG AGCTTATGCA GCTCCCTGAG GAGACCATTG CTGCCATCGA
2341 CGTGTCTTC GAGAACATCC GCAAGTTCCA CGCCGCCAG AAGGAGGAGA AGCCCTCCA
2401 GGTGAGACC ATGCCCGGTG TTGTCTGCAG CGTTTCTCT CGTCCCCTCG AGGCCGTCGG
2461 CTGCTACATC CCCGGCGGT ACGGCCTTCT CCCCAGCACT GCCCTTATGC TGGGTGTTCC
2521 CGCCATGGTC GCCGGCTGCA ACAAGATTGT GTTCGCTCT CCTCCCCGCG CCGACGGAAC
2581 CATCACTCCC GAGATTGTCC ACGTCGCTCA CAAGGTTGGG GCCGAGTCCA TCGTGCTTGC
2641 CGGCAGGTGCC CAGGCCGTAG CTGCCATGGC CTACGGCACC GAGAGCATCA CCAAGTCGA
2701 CAAGATTCTC GGCCCCGGTA ACCAGTTCGT CACTGCTGCC AAGATGTTCG TCAGCAACGA
2761 CACCAACGCT GCCGTTGGTA TTGACATGCC CGCTGGCCCG TCCGAGGTGC TGGTCATCGC
2821 TGACAAGGAC GCCAACCCCC CGTTCGTTGC CTCGGATCTC CTGTCCCAGG CTGAGCACGG
2881 CGTTGACAGT CAGGTCACTCC TGATCGCTAT TGACCTCGAC GAGGAGCATC TTCAGGCTAT
2941 TGAGGACGAG GTTCACCGTC AGGCTACGGA GCTTCCTCGC GTCCAGATTG TCCGTGGCTC
3001 CATCGCCAC TCGATCACCG TGCAGGTCAA GACCGTCGAG GAGGCCATGG AGCTCAGCAA
3061 CAAGTACGCT CCTGAGCACT TGATCCTCCA GATCAAGGAG GCCGAGAAGG CTGTCGATCT
3121 TGTGAAAC GCCGGTAGTG TCTTCATTGG CGCCTGGACT CCTGAGTCCG TTGGCGATTA
3181 CTCTGCTGGT GTTAACCACT CGCTGCGTAA GTTACATATC ATAAATAGCC CCGCTTCACA
3241 GATTCTTCTG CTAACGTCAA GACACATAGC TACCTATGGC TTTGGCAAGC AGTACTCTGG

```

Figure 8 continued

3301 CGTCAATTC GCCTCGTTCG TCAAGCACAT TACCAGCTCC AACTTGACTG CCGAGGGTCT
3361 CAAAAACGTC GGCCAGGCTG TCATGCAGTT GGCTAAGGTT GAGGAGCTCG AGGCTCACAG
3421 AAGGGCGGTC AGCATCCGTC TTGAGCACAT GAGCAAGAGC AACTAAACGG AAATTCTTT
3481 CGAAGTAGCA AAAAAAAA AAAAAAACAA GAACAAAAGG ATGTAGTGGG TTGATGTATA
3541 TCTGGGTCA TTTGGGCACA TAGAGTAATG ATAACGAGTT TTGGACATTG TACTGTTCTG
3601 TACAGGCTGA AGATCAGTAC ATGAATCTGT TGGTAAGTGT GGAGACCAA ACGTCCCTG
3661 AGTTTTCTC CCTATTCCAG AGGTGCTCGT CCCTGGGTGT TTATTTCAT TATTACATCA
3721 ACCTTTTTT TTTTTTTTT TTTTCAGAT CATCGTACA TGAACGGGG AAGCACAGAC
3781 GATCGAAACG TGGATGTCA AATGTCGCTG CAGTGATGCT GCATTGCATG AAGGCCAT
3841 CTCAATATAC TTGCAGTCTT GCACGTTGCA TGTGAACCTT CCAAACAACC GAATAAAAGA
3901 CGGCGAAAAA TGAAGATAAA AAAAAACCAT AAAAAAAATC AGAGGGAGTG TGGGAAATGG
3961 TGTCCTTTAG CATTTCAGACC CCATAGCCGT GCACGCCGG GTACAGACAG GTTCATCGAT
4021 GTTGACATTG ACTGGGACAC CAGGTCTATC TATTTTATCT CCTGTCCTCT ACCATACATC
4081 GGGACATCGG ACATCTTGCT GTACCCCCCA CACCCACAAA GCCTTATAAA AGCGCCACAC
4141 CCGAGGAGGT TCGGTGGCC CCACGAACTC TGTGCCTCCC TGCCTGTTA CAGGGACCGA
4201 ACGCTGGAGA ATCTTACTAG TTTCTGACA TCCGGCCTAC CCGAGCAGGA AAAGGGACAG
4261 CTCATAGGCG AGGAGGGATT TGAAGATGGG AACATTTGG GTGATTGAG AGGAGGAAC
4321 AGGTACTGCA TCATGATAGT TCGGGGCAGC ATCTTGGCTG GGACATTGTT AATACCTCGA
4381 TATGATGAAG TAGGAGGGAG TTTTGCGTG TCTTGCGAA GTCCAGAGAT CTGTTTATT
4441 TTATTTTTA TGGATGTAGT GTATCAACAC CCAAGATTG GAGAATAGTA CTAGGATTG
4501 CATTTACAAG TGGAACTCTT GAGAATCGTT GTATATCCTT GTCTTCCTG GAATGTTAAC
4561 AATCCTACAG CGAGCGAGCG AGCGGTGGA TGCGCTGATC TGATAGCGC AATATACGGC
4621 CGCTTCTCC GGTGCGTAG TGTAAGCTCT GTGGGCATAG TACACTAAA AAACCCCTG
4681 ATTCATGAT CTGCGTCTA TTCATTCCGA GCTATTCAG TGGTCACATT TCGAGGAAGA
4741 AAGAAAGCAA CTAAGATTCA CAGCCATCCA TCCATCCATA TCCAAGAATA ATCCATTCCC
4801 ATGTTCCCTC CCCCCCACTA TGTATGTGAC CACACGCTGC TGTCAAGAATG CCAACGGTCT
4861 CAGGTACCT CGTCCGACTG TTTGGCATGG AGTTACATAC ACTACTAGTG TAGCCCCGGG
4921 CCAAGCTACC CGTCAAATC TATACATATC TATAACGGGT TTCAGGGTT TCGTTCGCTG
4981 TCAATCAAGT TTGAAACATC ACTGGGGCCG TTGGACGGTG TATTAGACCA TTGGCTCCCT
5041 CAGCTGTTG GCGGCTGGC GGCTGGGTCA AACGGCAATA ACGGGACTCG AGAGGGACGA
5101 GGAGAGTCGG TTGGCTGGCT GCAATACAAG CGTTCCCACC TAACCAACGA GTCCCGTTT
5161 CCATTTGTGT GCCTAACCAT CATCTAGGGA TGTCAGGGTT TGGCCGGATC AGGGTATGTT
5221 TGGTTGACTG TTGTCATGTC TGATTGGTA CATATTATGG TAGGTGTCTC GAGAACAGTA
5281 GAGTACTCGG GCCTAGCGTT TGGATGATTA CGCGAGATAT GAGTTGTGGG CCGCCATGCA
5341 GTTGCTTGTC CATAAGCAGA AGTTGCTTT GGATATATT CTCGTCTTTC AAAGGTCACG
5401 AGGTCTGGG ACGAACGGCA TCGCCATCCA AAGGGTTGAA CATGAGAAAC CTGAATGGCC
5461 TTTGCGTTGA AATACAAAAA GTCAAGAACAA AAATCGCTTG AGGATAGGGA CGTGGAAAGCA
5521 AGCAAATATG GTAAGAGAGG TATACATCAA CCCTGGTTCA ATTGTTAGCG TGGTTCTTCC
5581 TCCACGTCT CGTTCATGAC GTTAAACAGT ACCAGGCTAA CAATTAAACC AGGGTTGATG
5641 TGTACTGATA TGTTAGGTGCT CAGCAAACG CCAATTCTT TGGCCCAAG CAGCAGTTG
5701 CTGTCAGTGC TGTCGTGTC AGCCTGGTA GTGGAACCTA AACTGCTAAC ACAGCGCAAG
5761 TGGCATGTA AAGATATTGT GGGAGGATCT GTATGGATGG ATGAGATTAC TGCTTGGTGT
5821 TGGTTGCGAG GCACTGCGGC TGTTAGGCTT TGCTGTGCC CGTTGACGA AGAAATACGC
5881 GGAACATATAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC GACGTAGCCG
5941 ACGGATTCTA GCAACATCCC GACTTTGCTT GTAGTGTACT ATGATAGCAG CACAGTGTG
6001 CTCCCTGTGA GAATGGGCTC TTTTTTTTT TCCCCCTTCC CTAGGGCGTT GACTGGACTT
6061 GCTCTATTGT TCCCAAGGTA GGTGCCGTC ATCGATTTC CCAAGTCTCC CGCCAGATTG
6121 TCGTCATAGT GTCATGATGA CCTCGGTGCG TGGGGCTGCG TGGTTACGGG GAGCTGGAC
6181 CGCTAGGCCT CAGTGGTTGT GCCATTCAAGC GTGGGTGTGT GGAGTAGCGG TAGAGGGCCT
6241 TGGAAAGTTGT GCTAGCGAA ACCCTGGAAT ATCTTCTACC CTCGATTCTCT TCTCGGGCTG
6301 CCCATGTGCT GAGGTGATGC CGGGGATCTG GCGCCAATCA TCCATTGAGG TTCCCGCAGC
6361 TTCCCGGTGC CGCGCGCGGG CGCAGTTGCT CACAGGACAC ACCTAGACGC AGGGGACAG
6421 GGGCACCGTT TGGTGTGCAA CTGGGTACCT AGCTGTAGCA AGCACTCCAC CGTCTGTGCA
6481 ATCCCCCAAT CCACGGCAGG AACTTCGCAC CGCCGCGGC CCGAGTGAGC GAATCCATCC
6541 GCATTGGATC CCAATTCTTGC CCCTTGCCAT CCTTCTTCT TCCCACCTGG CGCAACCAAC
6601 ACTTCCCTTG GTCTGGGTAC TCGTGTGAT CTTCACTCT TTTTTTCTT GGGCGACCGA
6661 CTTTTTATAT CCGTCCTTGC TTCCCCCTGG CCGTTGTCGT TCTTCTACA ACTACCTTCC
6721 GTTCATTATC CCCTTTCTTGC GTTCGGTCGA GGACCCAAAA ACAGAACAAAT TCCGGCTCTT
6781 CCAGGTGGCT TGGGTGCGAC TGTTTAGCTC TTGACCACTA GCCGCTTACCG TTCTCTTGAT
6841 GTTTTTATTT GGATATCATT AACTACTCT TTCTTGAAAC GGCAGACGAA CGGAACAGTT
6901 CCTACGGTAT ATTAGCGATA TACGTTGTAC TGATATTCTG AGCAAGAAGA GGCAAATTAT
6961 CAATTATGCA TCTCCCTTCG TCGCTGCTCA TCGCAGCTCC CTTGCTCGCC AATGTATCGG
7021 CCGAACCCAT TAGGATACCC CAACCGCAGT TTCTCCGTGG TATCAACATC ACAGCAACTT
7081 GCCGTTCGAG CACTACCGGA TTCGCCCAGC GGTGGATATG CCCCTGCCGT TGTAGACTGT
7141 CCAAGACCA AGCCGACGCT CCGGAAGGCC GTGGATTGT CGAACGGAGGA GAAGAACTGG
7201 TTGTCGATCC GGAGGAAGAA CACCATCCAG CCCATGAGGG ACCTCCTGAA GAGGGCCAAC

Figure 8 continued

7261 ATCACTGGGT TCGATTCCGA GACATTATG AATGAGGCCG CCAACAACAT CTCGCAACTG
7321 CCCAATGTCG CCATTGCCAT TTCAGGAGGC GGCTATCGT CCCTCATGAA CGCGGCCGGC
7381 TTCGTTGCTG CTGCGGATAA CCGAATTCAA AATACCACGG GCGCAGGTGG TATTGGAGGC
7441 TTGTTGCAGT CCAGCACATA TTTGTATGTA AAGTGGTTCT TCTTATCTCG TTTTCGAGTG
7501 TCAACTGCGC CAGTCAGAG TTGGCGGGCT GTGGACGACC TTGCTGGTGA ACATGTCTTG
7561 GACTCCATGC CCCTTCTCG TTTCTCAA TCAAGAACGT GAGGACCGTG ACCGTAATC
7621 GCTAACGCAA CTCTAGGGCC GGACTTTCTG GTGGTGGCTG GCTTGTGGC AGTTTGTCT
7681 CCAACAACTT CAGCAGCATT GAGACCCTGC TGAGCGAGAA CAAAGTCTGG GACTTTGAGA
7741 ACTCCATCTT TAAAGGGCCC AAGGAGGCTG GCCTTAGTAC TGTCAACCAC ATTCAAGTACT
7801 GGTCCGAAGT GGCAAAGGAA GTTGCXAAGA AGAAGGATGC TGGCTTCGAG ACAAGTATAA
7861 CAGACTACTG GGGCGAGCA TTGAGTTACC AACTGATCGG AGCCGATATG GGCGGCCCGG
7921 CTTACACCTT CTCCAGCATT GCCCAGACCG ACAACTCCA GAAGGCCGAA ACGCCGTTCC
7981 CTATTCTGGT AGCTGACGGC CGCGCGCTG GAGACACCAC CATCTCCCTC AATGCTACCA
8041 ACTACGAGTT CAACCCGTTG GAGACGGTA GCTGGGACCC GACCGTCTAT GGCTTGCAGC
8101 CGACCAAGTA CCTCGGCGCC AACTTCAGCA ACGGCGTGAT CCCATGGGA GGCAAGTGC
8161 TTGAGGGTCT CGACCAAGCC GGCTTCGTCA TGGGCACCAG CAGCACGCTC TTCAACCAGT
8221 TCCTTTGGC CAACATCTCC AGCTACGACG GTGTTGCCCG ACCTGCTCAT CGAAGCCGTG
8281 ACTTCTGTCC TCAAGGAAAT CGGCGCCAAG AGGACGACGT CTCCCAAATC ATCCCTAATC
8341 CGTTCTGGA CTGGAACAAAC CGGACCAACC CCAACGCCGA CACGCTCGAG CTCGACCTGG
8401 TCGACGGCGG CGAAGATCTG CAGAATATTC CGCTCAACCC GCTCACCCAA CCCGTGCG
8461 CCGTGGACGT CATCTTCGCT GTGCACTCGT CCGCCGACGT GACAAACTGG CCCAATGGCA
8521 CCGCCCTGCG AGCCACCTAC GAGCGCACTT TCGGCTCTAT TTCCAACGGG AACTCTTCC
8581 CCTCGATCCC CGACGACTGG ACGTTTATAA ACCTAGGCCT CAACAACCGC CCCTCTTCT
8641 TCGGCTGCGA TGTAAAGAAC TTTACCTTGA ACGCCAACCA AAAGGTTCCC CCCTTAATCG
8701 TCTATGTCCC CAACCGGCC TATACCGCGC TGAGCAACGT GTCCACCTTC GATCCGTCT
8761 ACACCATGTC TCAGCGCAAC GACATCATCG GCAACGGATG GAACTCAGCC ACGCAGGGAA
8821 ACGGCACGCT GGATTGGAG TGGCCCACCTT GCGTCGCGCTG CGCGGTTATC AGCAGGAGCT
8881 TAGATCGGTT GGGCAGGCAG ACGCCAGCCG CGTGCAAGAC TTGCTTGAG AGGTATTGCT
8941 GGAATGGCAC AGTGAACCTCA AAAGATACAG GGGTTTACAT GCCTGAGTTC AAGATTGCG
9001 ATGCGCATGC CCTGGACTCG GGTGCTGTTG CTATCGAAA GATGGTGAAT GTCTGGTCGT
9061 CGGTTGTGGT GGGAGTTGTG GCGGCTACTT TGTTGTTGTA GGGGTAGGGG AGACGTGATG
9121 ATATTCCAGT CTGATGAAGT TGAGACTGG A CTGGAGATCG CCAAGGATGC GGAGGGAAAG
9181 GAATGCGTGG TGTAAATGTC ATGATGGATG AAGGGTCATG GATCATGGAA CGACGGGGCG
9241 GGGATATTGG ATGATGGATA TACCAACTG CATGCATGCT CTATTGATAA TATGCTTTGG
9301 CATTACGTT TAACAATCAA TTGCTCCATC CTGATGTTCT ATCTTCGAC ACTGGATTGA
9361 TACTACTCCT GTGCTTCCC TCTTGAAGTT GGAAGGACTT GAGGTTGGAA GGACTTGAGG
9421 TTGTTGTTTC TGAGGGAGGT TATCGAAGTA TCATCTGTGC TGATGCCGAT CGATAGACTG
9481 CCCTCTTCTT CGAGGCAACG AACGGTCGGA TGAGCCTCTA ATCATGATGC TCAGTGCCAC
9541 AAAAAGGCTC CAGCACAGCT GCCCACACCT TTTTGCGCTC GTGCTCCTT CCTTTTTTC
9601 CCCCCCTTTC TTCTTCCA TCTCATCCCG TACCAAGGTG CCCACCGGGT ATATATATTA
9661 CCTCCTTGGC CGTTCTCCTT TGACCAATAA ATCGCTTGGT CGAGTGGCGT AACCGTTTAC
9721 CGTCTACACT TATCACTCAA ACCAAACCAA ACCATCGAAG AAGTTACCTA TCGGTTCGAG
9781 GGAACGGTGA TGTCTTACG TTCAAGTTAA CCCAAAGAGC GTTCCACATC GTTGAACCGT
9841 CTCCTCCAGT TCTTGGATCT GTTTAACCTC CGCAGCGACT GAAGAAGTAA TCACTTTTT
9901 TTTTTTGTT TCCAAAAAAA AAAAAAAA TTAC

Figure 9 Construction of the components of the sequence diversification cross: Parent (variant 1) and Parent (variant 2). For convenience, plasmid sequences are shown as linear. The cross hatched region in the chromosome is dispensable. Stippled sequences in the plasmid indicate the multiple cloning site for inserting foreign DNA. Crossovers in region 1 and region 2 insert the foreign sequence to be diversified into chromosome 1 of *Neurospora crassa* adjacent to the recombination hotspot *cog*. Parent (variant 2) containing a version of the foreign sequence with multiple differences from that in parent (variant 1) is similarly constructed. Parent (variant 1) and parent (variant 2) are crossed and conversion events (stippled arrow) initiated (X) at *cog^L* recombine the sequence differences in variant 1 and variant 2 to form new combinations. Sequences are identical except for those that distinguish variant 1 and variant 2. *rec-2* on linkage group V permits *cog^L* to be active. For simplicity, genes not directly related to the diversification are omitted. See text for further details.

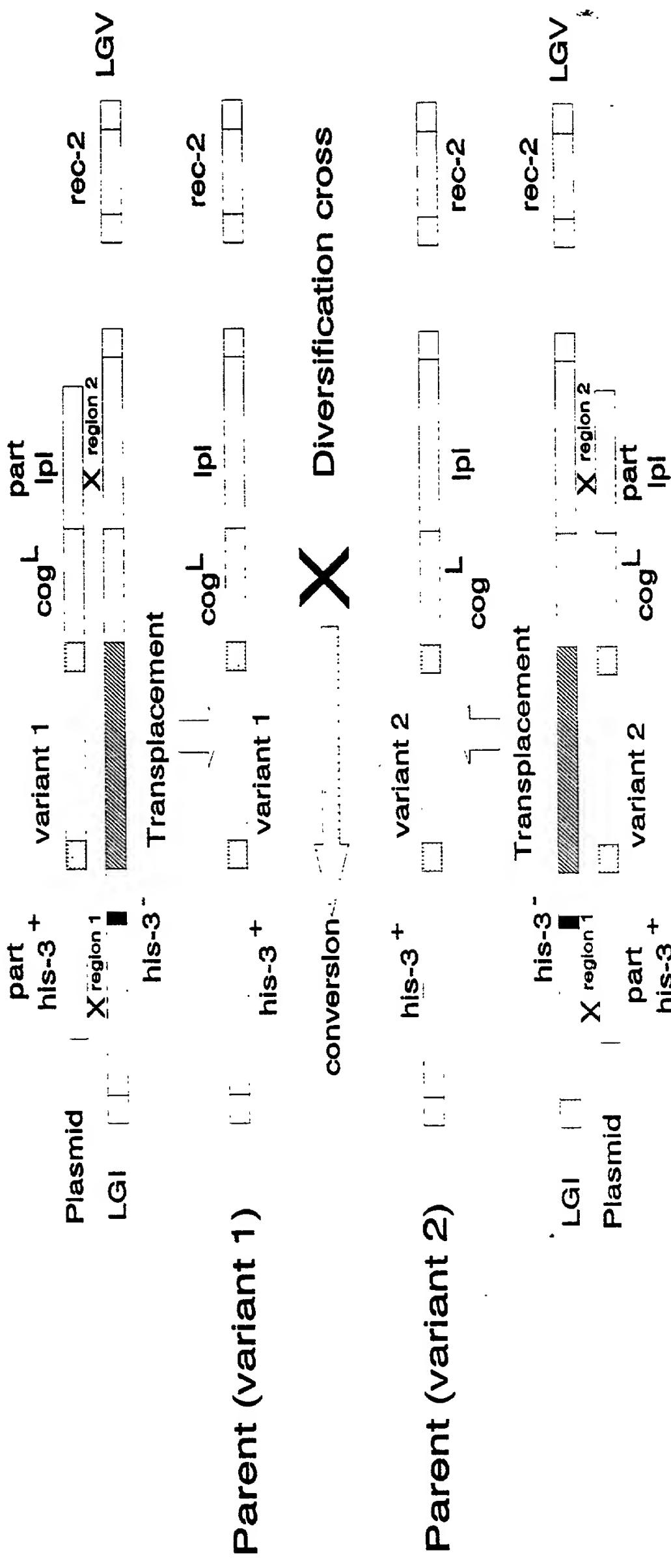


Figure 10 Construction of parent (variant 1) and parent (variant 2) enabling selection of progeny that have experienced conversion in the foreign DNA. Complementing pairs of *his-3* alleles are used to obtain parent (variant 1) and a different pair of complementing *his-3* alleles are used to obtain parent (variant 2) as explained in the text. Parent (variant 1) and parent (variant 2) are crossed and *his⁺* recombinants are selected. These must all have experienced conversion events affecting the foreign DNA since the events begin at *cog^L*. The *his-3* alleles in parent (variant 1) and parent (variant 2) are non complementing to ensure that selection yields recombinants and not aneuploid progeny having two copies of all or part of linkage group I.

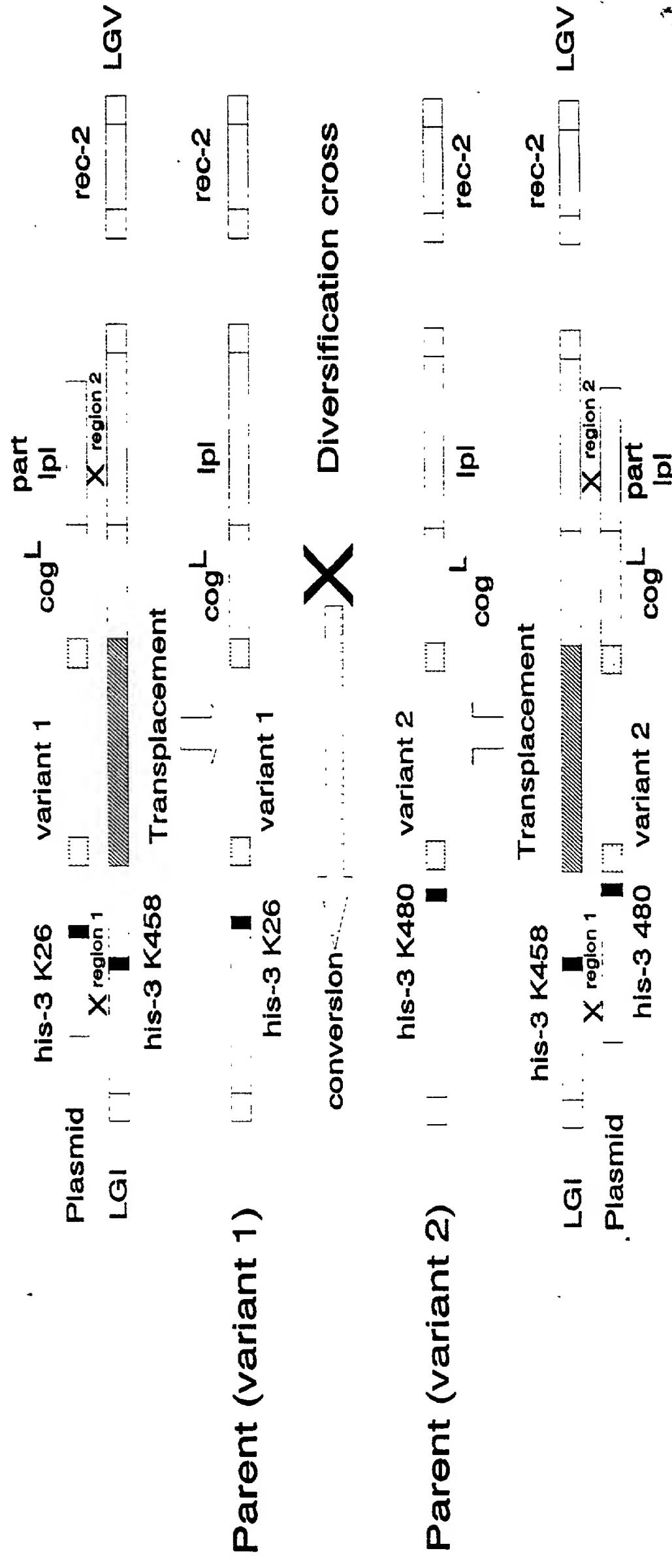
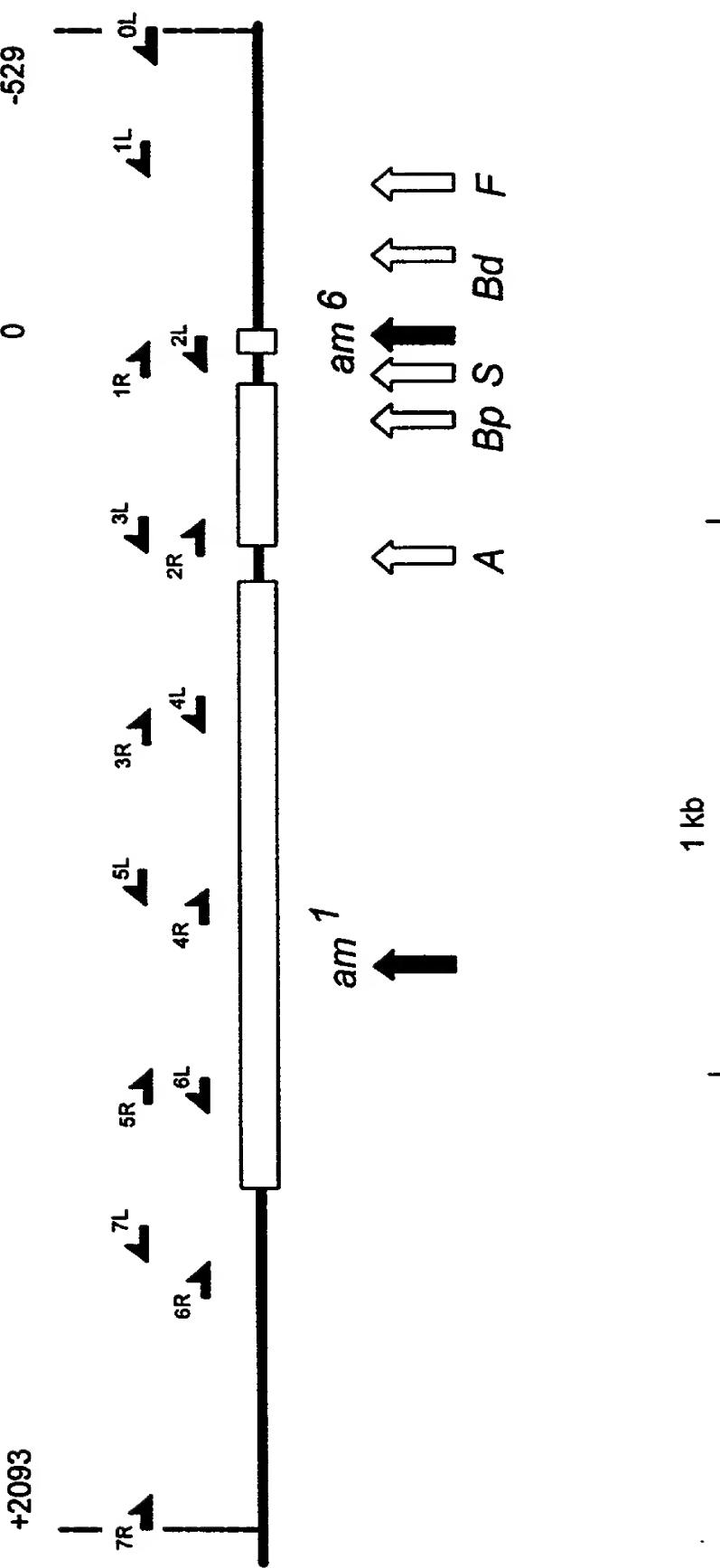


Fig 11



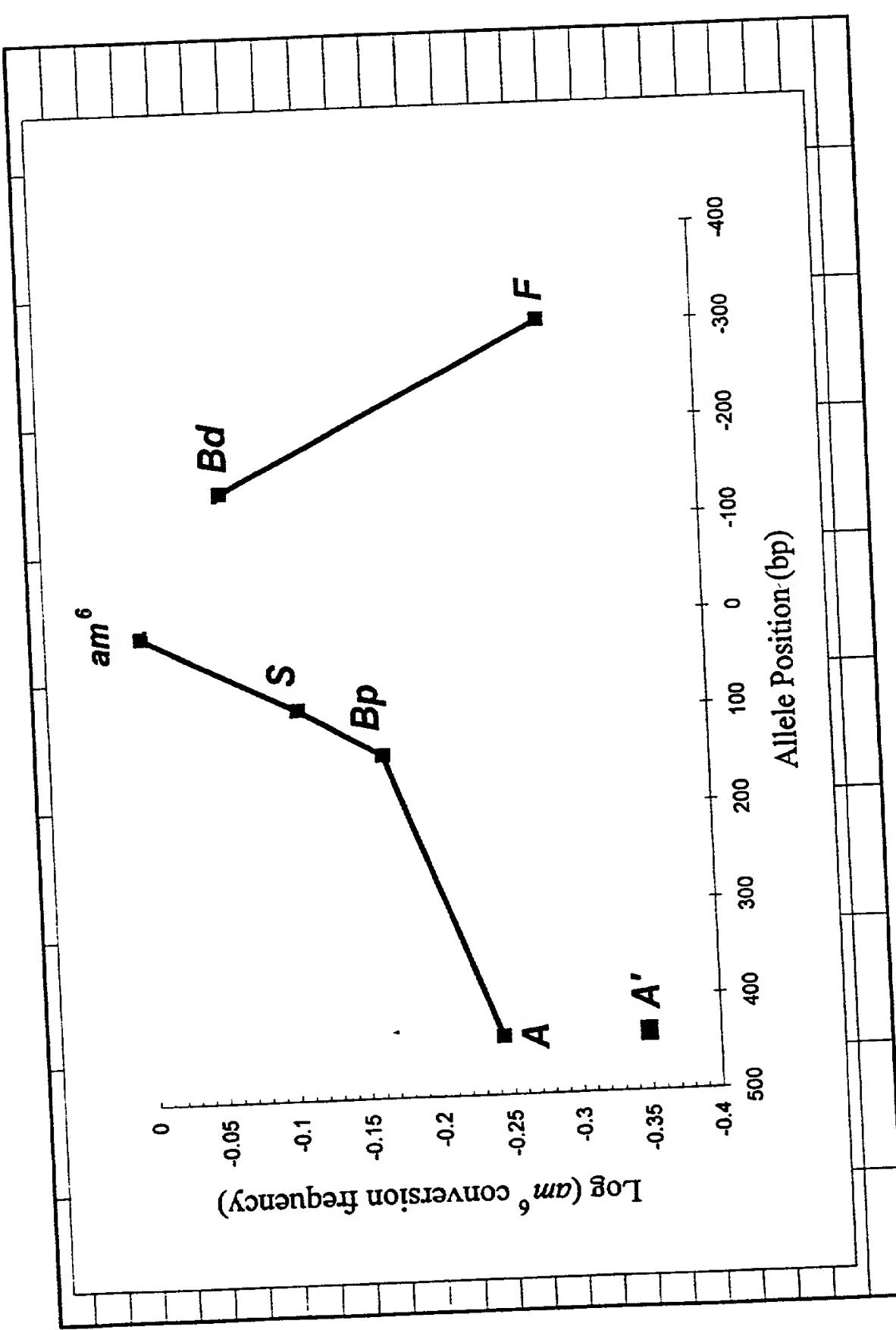


Fig 12

Fig 13

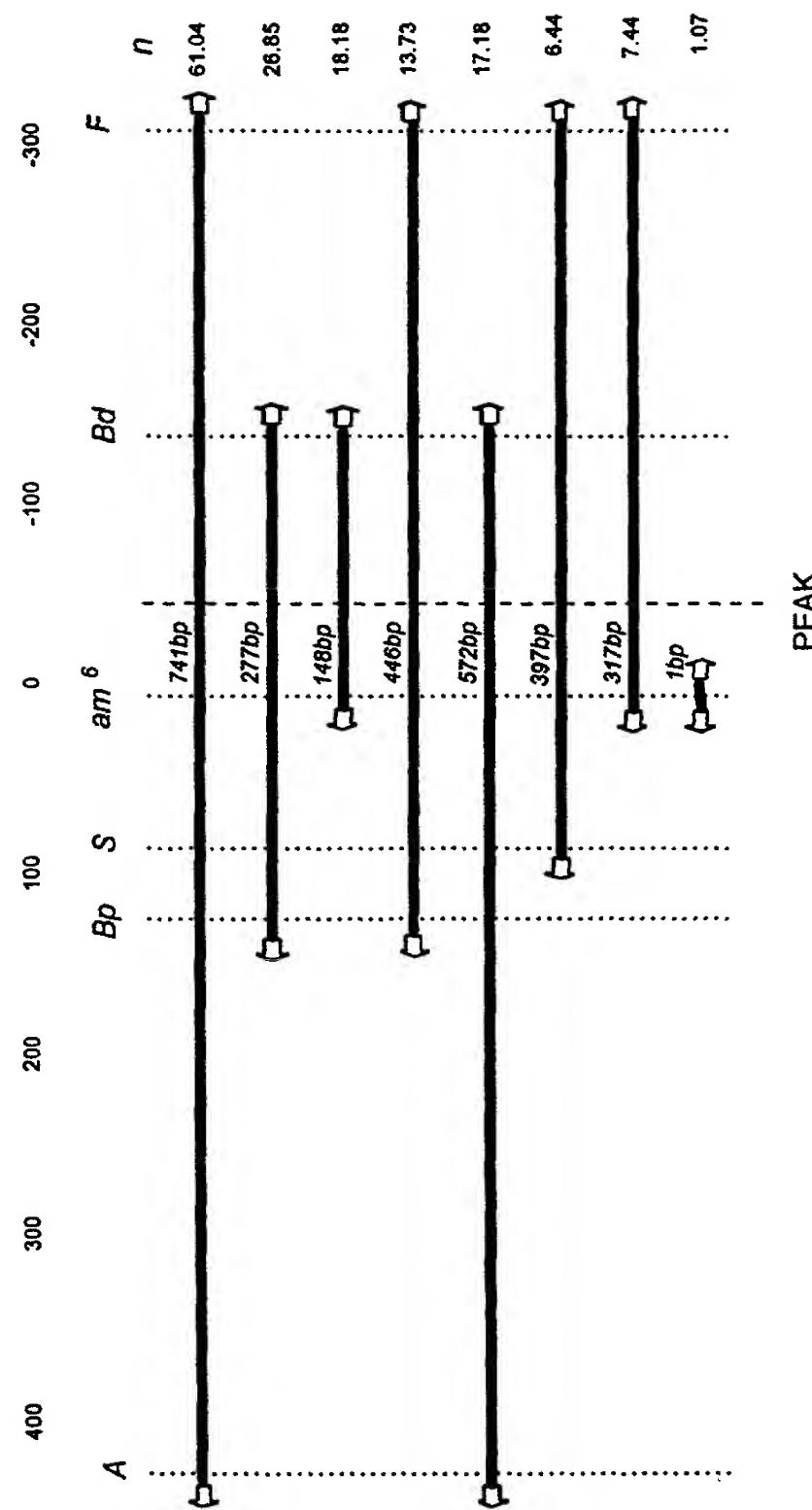
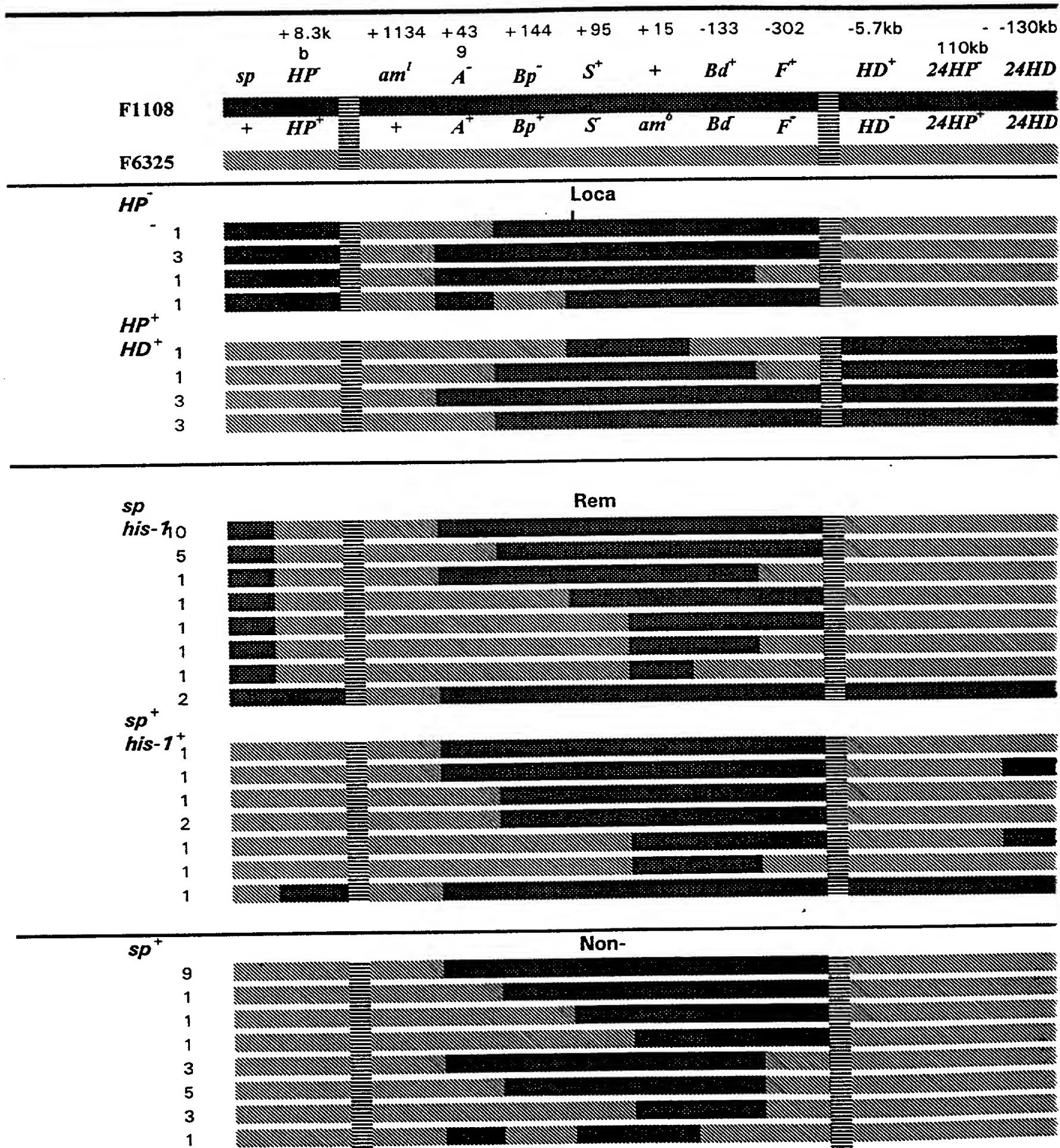


Fig 14
(2 pages)



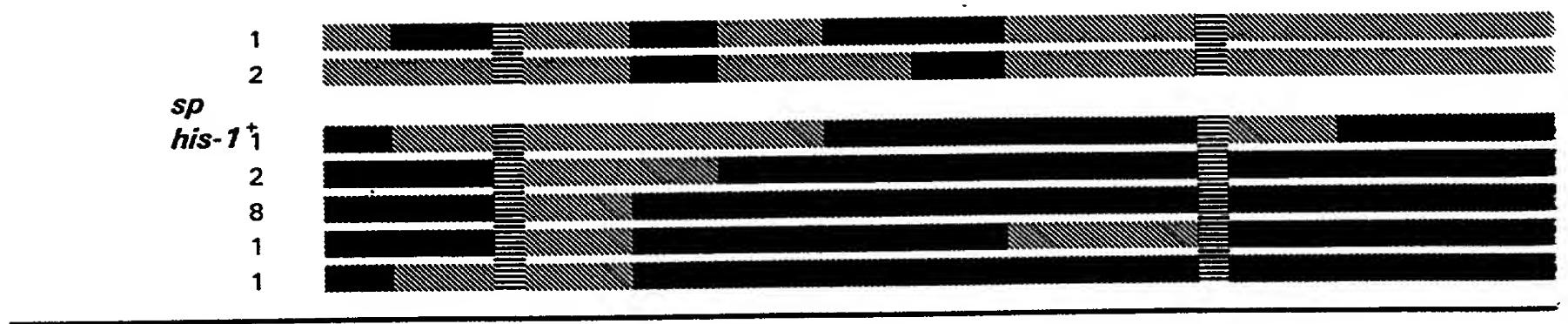


Fig. 14 (continued)

FIG. 1A

Methods for the diversification of DNA sequences and testing for superior variants

existing protocols: *Number of transfections needed to generate 1024 new variants: 1024*

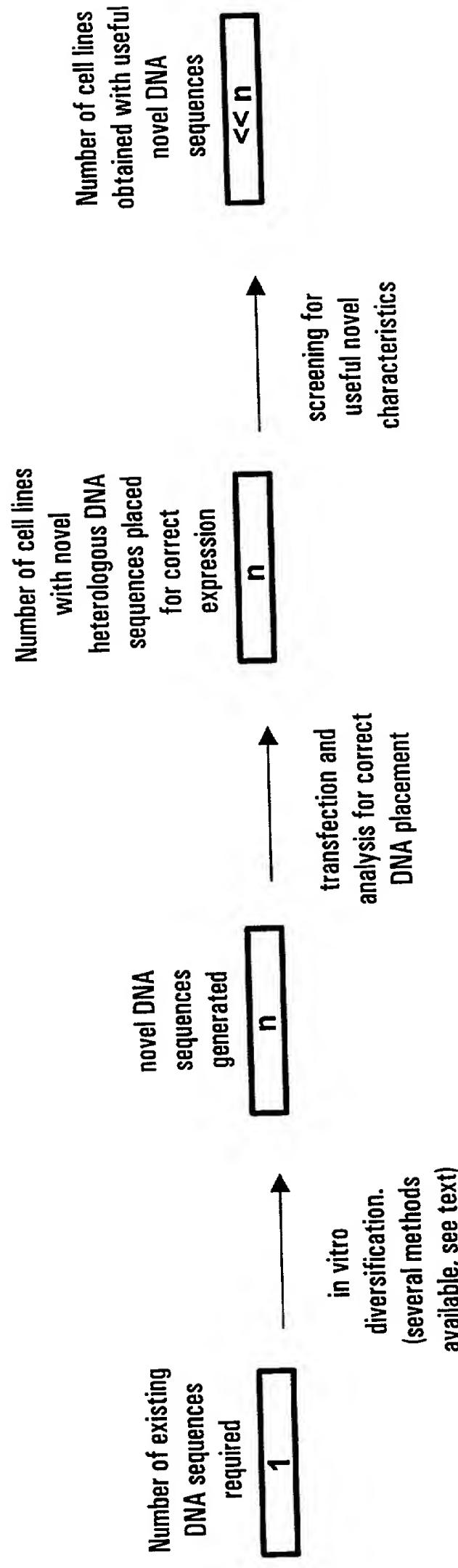


FIG. 1B

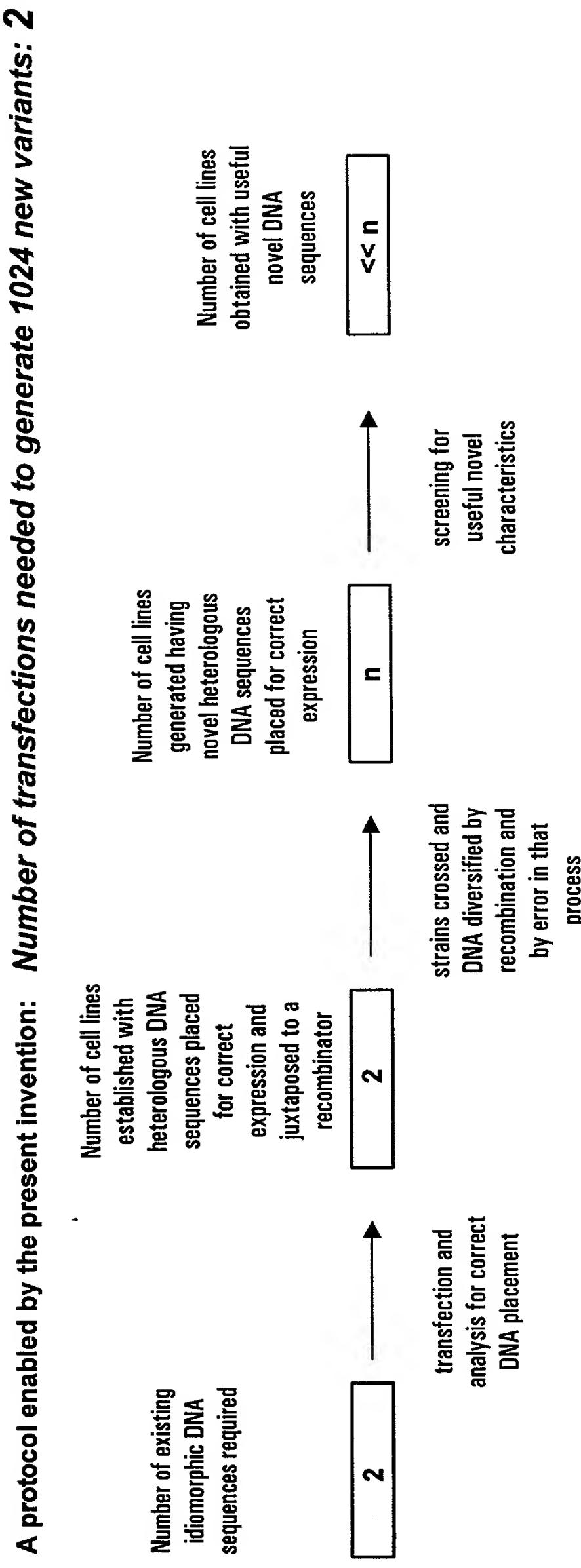


FIG. 2A

Methods for the diversification of DNA sequences coding subunits of heteropolymeric proteins and testing for superior variants.

The example given for immunoglobulins is for illustrative purposes only and is not intended to limit application of the present invention to this specific heteromeric protein. H = heavy chain genes, L = light chain genes

Existing protocol: Number of transfections needed to generate 1024 new combinations: 2048

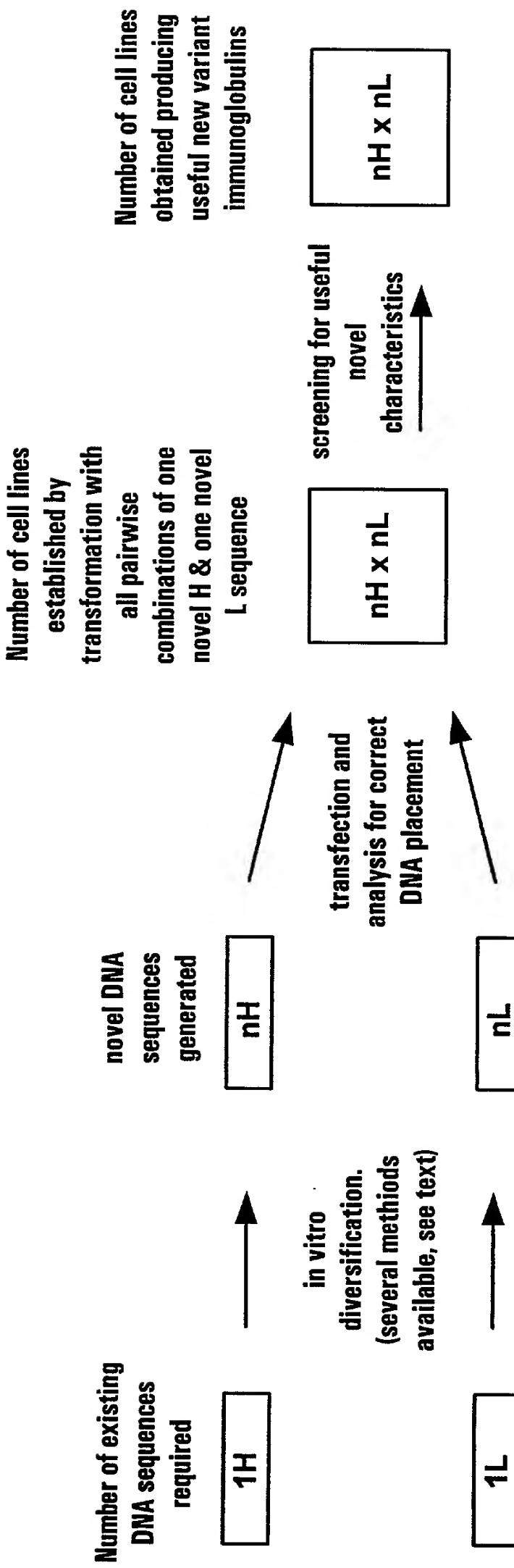


FIG. 2B

Existing protocol using the heterokaryon technology of US Patent Serial No. 5,643,745
Number of transfections needed to generate 1024 new combinations: 64

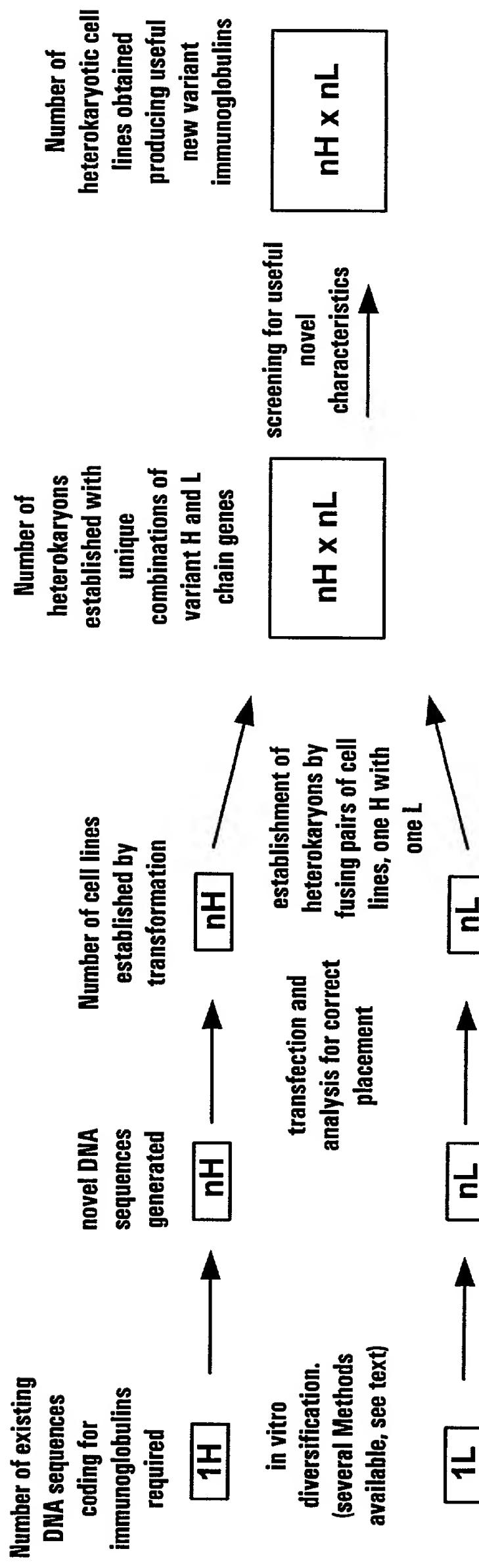


FIG. 2C

A protocol enabled by the present invention

Number of transfusions needed to generate 1024 new combinations: 4

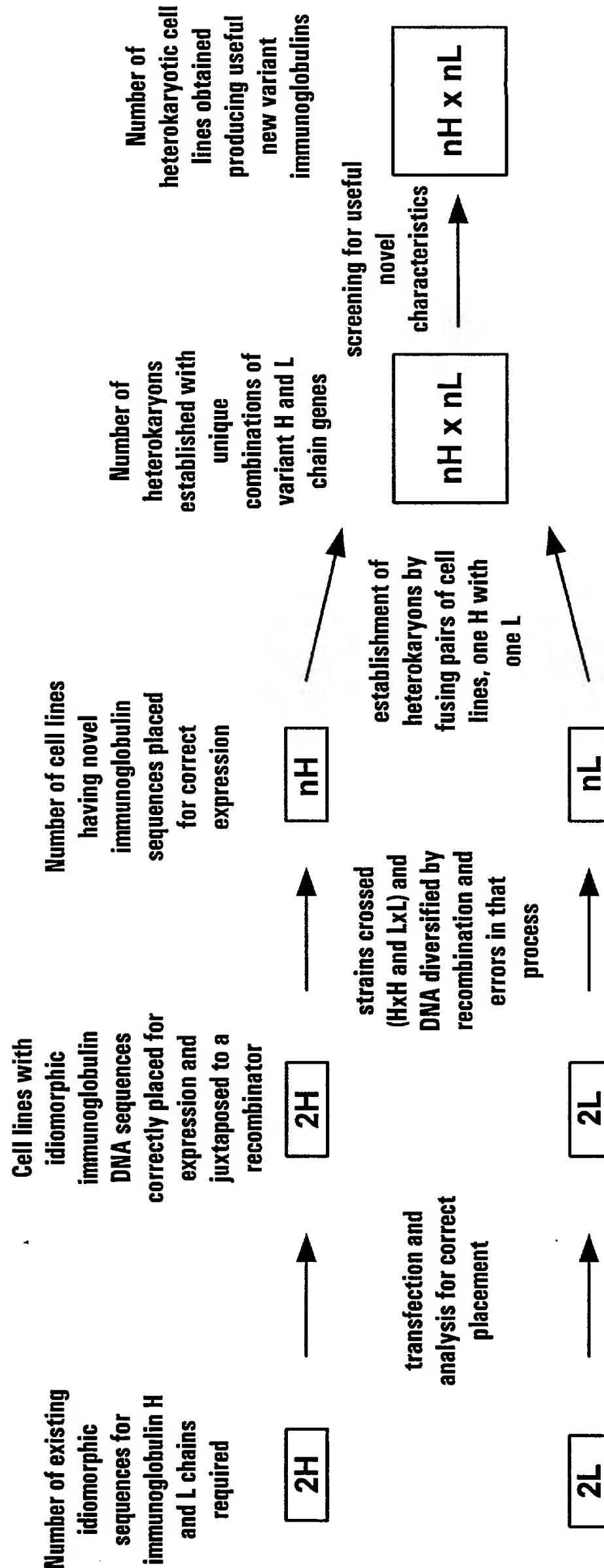


Fig. 3

The modified double strand break repair model for meiotic recombination. After H Sunet *et al* Cell 64: 1155-1161, 1991

(a) A double strand break (DSB) is made in one DNA duplex. (b) A long 3' overhanging single strand tail is generated either side of the break by resection. (c) One 3' end invades a homologous duplex forming a D loop. (d) the D loop is enlarged by repair synthesis and anneals to the second 3' end (e) Repair synthesis occurs at the second 3' end and two intermolecular junctions (Holliday junctions) are formed. Resolution of the junctions by cutting inner and outer strands can give rise to non-crossover (f) and crossover (g) chromosomes. If there are base mismatches in the heteroduplex regions (duplex molecules with thick and thin lines) there will be gene conversion. If mismatch repair does not occur there will be post meiotic segregation of new sequence combinations.

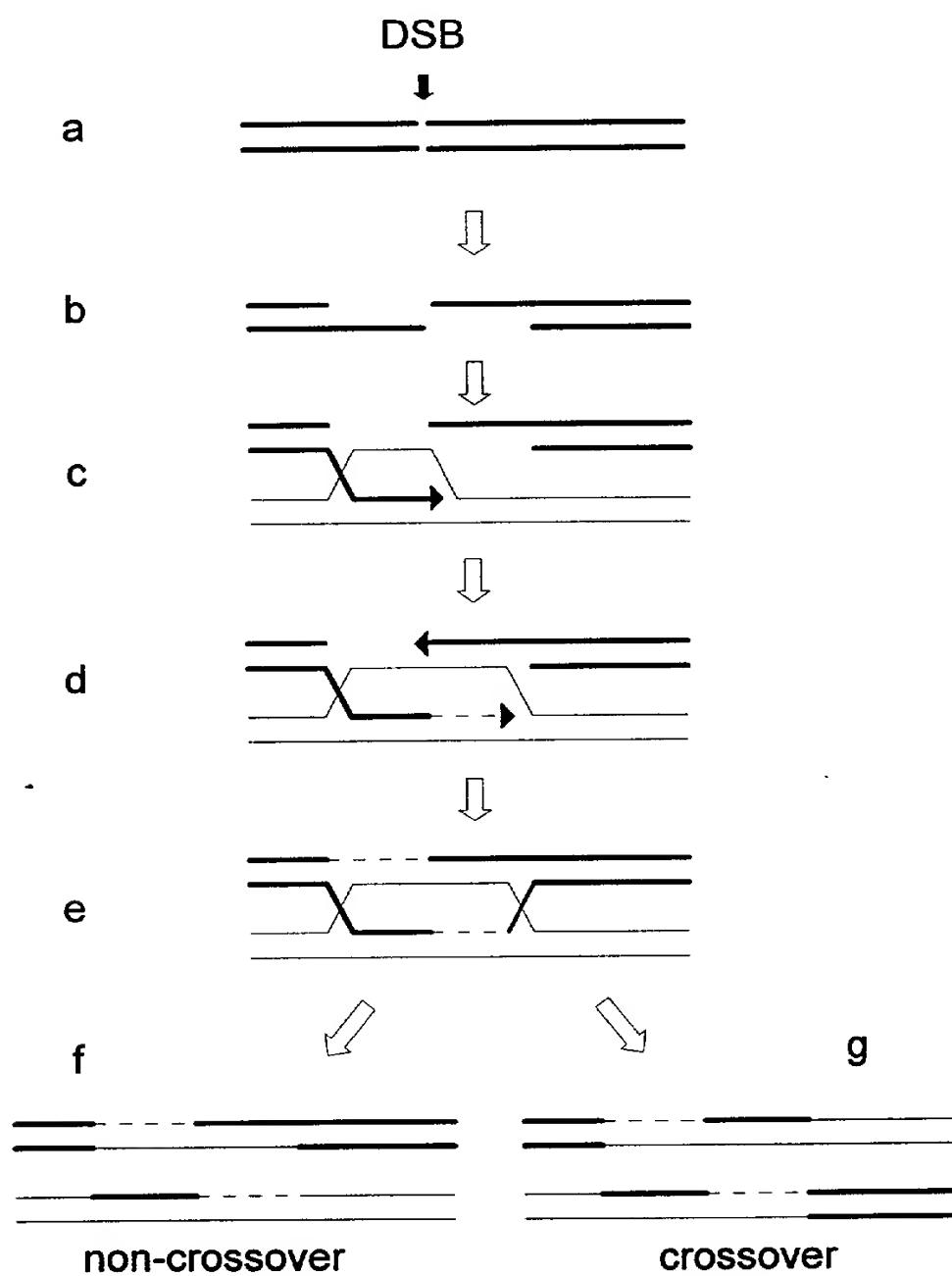


FIG. 4

Life cycle of *Neurospora crassa* after JRS Fincham (Genetics, Wright 1983). Microconidia having one nucleus are not shown but can be generated as described in the text. Perithecia and protoperithecia are shown in section.

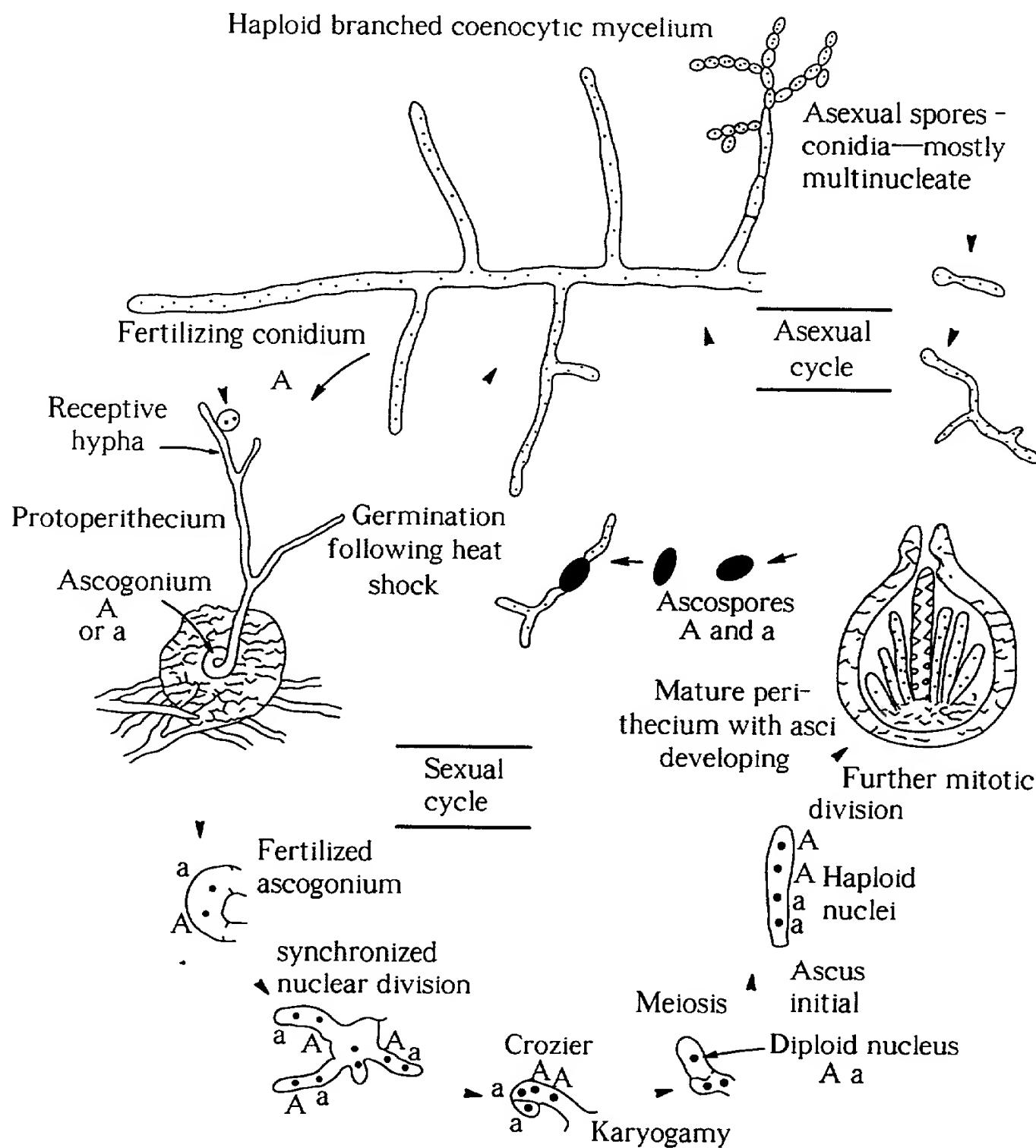


FIG. 5

Map of the *his-3*, *cog*, *lpl* region of Linkage Group I of *Neurospora crassa*. Vertical bars, triangles and hairpins show the location of sequence differences that distinguish the St Lawrence and Lindegren wild type strains. The corresponding full DNA sequences are given in Fig. 7 and Fig. 8. Vertical slashes indicate one to seven base substitutions per 10 base pairs. Triangles indicate short sequence insertions and the hairpin a 101 base pair inverted repeat transposon fragment present in St Lawrence.

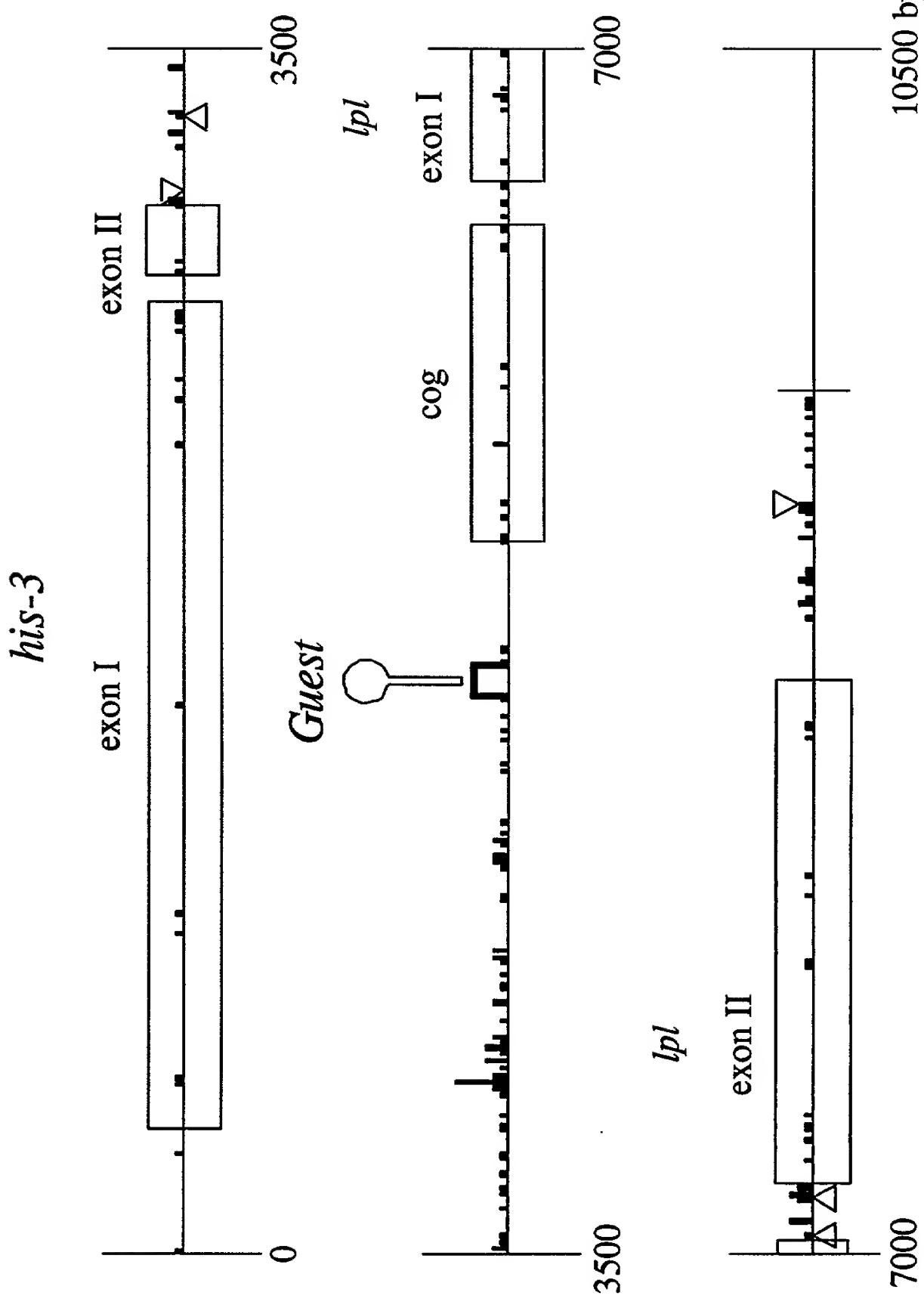


FIG. 6

Discontinuity in the parental origin of DNA sequences in progeny from crosses between pairs of *his-3* alleles. In most cases this reflects discontinuity of conversion tracts, in some cases crossovers near the ends of conversion tracts. Markers are specific DNA sequence differences that distinguish the parents. These were all E (Emerson wild type origin) or all L (Lindegren wild type origin) in the parental strains. Recombinants carry both E and L markers. Marker position is given in base pairs from the first base of the first codon of the *his-3* gene. Each line of the table shows the parental origin of the markers inherited by one of the progeny.

marker location	P	H	P1	K1201	K504	L3	R1	K26	K874	R4	C4	C5	C6	C6/7	C7	C8	C1	C2	C3	C9	D
~6000 -384 115 179 563 1232 1502 1717 2318 3436 3705 4000 4304 4667 4821 5232 5495 6153 6507																					

his-3

cog

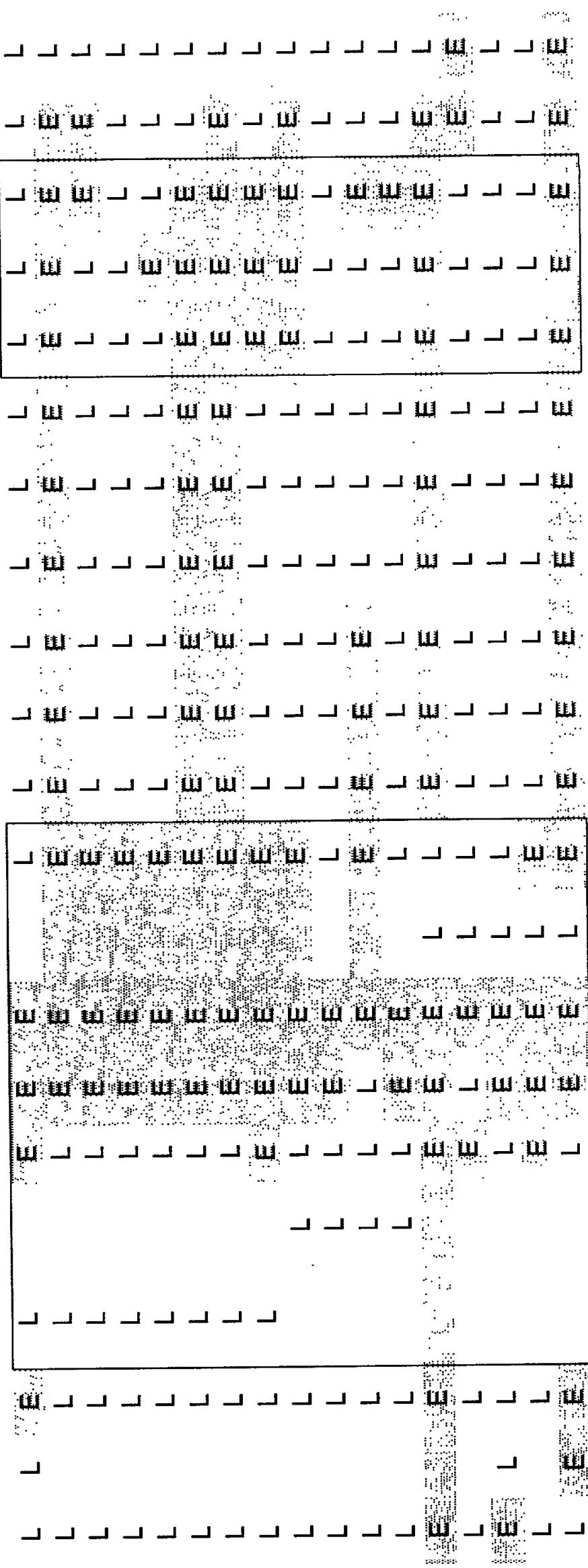


FIG. 7

Nucleotide sequence of the *his-3 cog^L lpl* region of linkage group I in the Lindegren wild type strain of *Neurospora crassa*. This differs from that in the StLawrence strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the high frequency recombinator *cog^L* which is active providing the cross in which meiosis occurs is homozygous *rec-2*.

1 GATCGCAACT GGAGATCACT CGCACCGTGC CGCAGAACAA GGGCGACGAG CCTCAGGGCA
61 GTTAGCCTG CCGTAACAGC ACAGACCATA GCTTATTTC ACCTGGCCGG GCGGGCGACG
121 GCGGCACTGA CATCGGCAAG GCGGCATCAA GCAACCCCTC TGGTGCTTC CAGCTGCCGG
181 CCAACGTCAAG CGGTACAAGG AGAAATCTGG AAGGAAAGAC TTCTGGCACC GACAGGATGG
241 CACGCGGGAA AAGTTCCCAA TGCATGAGAT GAGGGGCATT TGCAATTGCCT CCCGTCACAC
301 TGCCCGCGAA CCCCAACCCC ACCATAGCGT CTGTCGATAC ATGGAGCGCG AAGTCGAGAA
361 ACCTGTAATT CCTGGTAACT TTCAGGTACA CAGTACGTAC TGATCCTGGT ATCAAACCTT
421 GCCTGCCGAG TTTTCGACGG AAAGAGGTGT GAATTGTGAA AGAGTCATAC CAAATCACCC
481 GATTTTCATA AAGCCCGAGT CTTTCTGTA CATAAGCGAC ACTCGAAGCG GGCCTCATCT
541 TCATAGCCTG ATAGCTTGTAT ATACTCCATC CTCGTATCTC ACTTGACCTT GAGTTCAACC
601 CCACGTCAGA CTTCACCCGA CACATCGACG GATTGGGGAA CAGCACAATA CCTGAAAAGC
661 GAGAAAACCA AACAGAGGAA AACACCATGG AGACAACACT TCCCCCTCCCC TTCTCGTCG
721 GTGTCAGTGT TCCTCCCGGA CTGAATGACA TCAAGGAGGG CCTCAGCCGG GAGGAAGTCT
781 CGTGTCTTGG CTGCGTCTTC TTCGAGGTCA AGCCAAGAC CCTTGAGAAA ATCGTGCAG
841 TCCTCAAGCG TCACAATGTC GAATTGAGC CCTACTTCGA TGTAACAGCC CTCGAGTCTA
901 TCGATGATAT TATCACTCTT CTGGACGCCG GCGCCCGCAA GGTGTTGTC AAGACCGAGC
961 AGTTGGCCGA CCTCTCCGCA TATGGCTCCC GCGTTGCCCG CATTGTCACT GGAAGCAGCG
1021 CTGCTTGCT TTCCCTCCGCC ACCGAGAGCG GCCTTTGCT CTCCGGCTTC GATCAGACTG
1081 CCTCCGAGGC TGCACAGTTT CTGGAGGAGG CCAGAGACAA GAAAATTACC CCCTTCTTCA
1141 TCAAGCCCGT TCCTGGGGCC GATCTCGAAC AGTTCATCCA GGTGCGCCGCC AAGGCTAACG
1201 CCATCCCCAT CCTGCCATCC ACTGGCTTGA CAACAAAGAA GGACGAGGCC GGAAAGCTTG
1261 CCATCTCCAC CATCCTCTCG AGCGTCTGGA AGTCTGACCG TCCCGATGGT CTGCTCCCCA
1321 CCGTTGTCGT TGATGAGCAC GACACTGCTC TGGGTCTGGT CTACAGCAGT GCCGAGAGTG
1381 TGAACGAGGC CCTCAGGACA CAGACTGGTG TCTATCAGAG CCGGAAGCGC GGTCTCTGGT
1441 ACAAGGGTGC TACTTCCCGA GACACTCAGG AGCTCGTCCG CATCTCGCTT GACTGCGATA
1501 ACGATGCTCT CAAGTTGTC GTGAAGCAGA AGGGTCGTTT CTGCCACCTC GATCAGTCCG
1561 GCTGCTTTGG TCAGCTAAA GCCCTCCCA AGCTCGAGCA GACTTGATT TCGAGGAAAC
1621 AGTCTGCCCG CGAGGGCTCC TACACTGCCG GTCTCTTCTC CGATGAGAAG CTAGTCCGGG
1681 CCAAGATCAT GGAGGAGGCT GAGGAGCTCT GCACCGCTCA GACCCCCCAG GAAATCGCCT
1741 TTGAGGCTGC CGATCTCTTC TACTTGCTC TTACCAAGGGC CGTTGCTGCC GGCAGTTACTC
1801 TTGCCGATAT CGAAAGGAGC CTTGACGCCA AGAGCTGGAA GGTCAAGCGC AGGACTGGAG
1861 ATGCTAAGGG TAAGTGGGCT GAGAAGGAGG GCATCAAGCC TGCGCGTCC GCTCCCGCTG
1921 CCACTTCGGC CCCTGTCACC AAGGAGGCCG CCCAGGAGAC CACCCCTGAG AAGATCACCA
1981 TGAGACGTTT CGACGCCTCC AAGGTCTCTA CCGAGGAGCT CGATGCTGCT CTCAAGCGTC
2041 CTGCGAAAAA GTCGTCCGAT GCCATCTACA AGATCATTGT CCCCCATCATC GAGGACGTCC
2101 GCAAGAACGG CGACAAGGCT GTTCTGTCGT ACACTCACAA GTTCGAGAAG GCTACCTCTC
2161 TTACTAGCCC CGTCCTGAAG GCGCCCTTCC CCAAGGAGCT TATGCAGCTC CCTGAGGAGA
2221 CCATTGCTGC CATCGACGTG TCCCTCGAGA ACATCCGCAA GTTCCACGCC GCCCAGAAGG
2281 AGGAGAACGCC CCTCCAGGTC GAGACCATGC CCGGTGTTGT CTGCAGCCGT TTCTCTCGTC
2341 CCATCGAGGC CGTCGGCTGC TACATCCCCG GCGGTACCGC CGTTCTCCCC AGCACTGCC
2401 TTATGCTGGG TGTCCCGCC ATGGTCGCCG GCTGCAACAA GATTGTGTT GCCTCTCCTC
2461 CCCGCGCCGA CGGAACCATC ACTCCCGAGA TTGTCCACGT CGCTCACAAG GTTGGGGCCG
2521 AGTCCATCGT GCTTGCCGGC GGTGCCAGG CCGTAGCTGC CATGGCCTAC GGCACCGAGA
2581 GCATCACCAA GGTGACAAG ATTCTCGGCC CCGGTAACCA GTTCGTCACT GCTGCCAAGA
2641 TGTTGTCAG CAACGACACC AACGCTGCCG TTGGGATTGA CATGCCCGCT GGCCCGTCCG
2701 AGGTGCTGGT CATCGCTGAC AAGGACGCCA ACCCCGCGTT CGTTGCCTCG GATCTCCTGT
2761 CCCAGGCTGA GCACGGCGTT GACAGTCAGG TCATCCTGAT CGCTATTAAC CTCGACGAGG

FIG. 7 continued

2821 AGCATCTTCA GGCTATTGAG GACGAGGTT ACCGTCAAGGC TATGGAGCTT CCTCGCGTCC
2881 AGATTGTCCG TGGCTCCATC GCCCACTCGA TCACCGTGCA GGTCAAGACC GTGCGAGGAGG
2941 CCATGGAGCT CAGCAACAAG TACGCTCCTG AGCACTTGAT CCTCCAGATC AAGGAGGCCG
3001 AGAAAGCTGT CGATCTTGTC ATGAACGCTG GTAGTGTCTT CATTGGCGCT TGGACTCCTG
3061 AGTCCGTTGG CGATTACTCT GCTGGTGTAA ACCACTCGCT GCGTAAGTTA CATATCATAA
3121 ATAGCCCCGC TTCACAGATT CTTCTGCTAA CGTCAAGACA CATAGCTACC TATGGTTTG
3181 GCAAGCAGTA CTCTGGCGTC AATCTCGCCT CGTCGTCAA GCACATTACC AGCTCCAAGT
3241 TGACTGCCGA GGGTCTCAAA AACGTGGGCC AGGCTGTCAT GCAGTTGGCT AAGGTTGAGG
3301 AGCTCGAGGC TCACAGAAGG GCGGTCAAGCA TCCGTCTTGA GCACATGAGC AAGAGCAACT
3361 AGACGGAAAT TCTTTTCGA AGTGCAAAA AAAACAAGAA CAAAAGGATG TAGTGGTTG
3421 ATGTATATCT GGGTCATTT GGGCACATAG AGTAATGATA ACGAGTTTG GACATTGTAC
3481 TGTTCTGTAC AGGCTGAAGA TCAGTACATG AATCTGTTGG TAAGTGTAGA GACCCAAACG
3541 TCCCTTGAGT TTTTCTCCCT GTTCCAGAGA GGTGCTCGTC CCTGGGTGTT TATTTTCATT
3601 ATTACATCAA CCTTTTATT TATTTTATT TTTTATTAC TTTTTTTCC TTTTTTCAG
3661 ATCATGCGTA CATGAACGGG GGAAGCACAG ACGATCGAAA CGTGGATGTC ACAATGTCGC
3721 TGCAGTGATG CTGCATTGCA TGAAGCGCCC ATCTCAATAT ACTTGCAGTC TTGCGCGTTG
3781 CACGTGAAC TCCCCAACAA CCGAATAAAA GACGGCGAAA AATGAAGATA AAAAAAAACC
3841 ATAATAAAAAA TCGGAGGGAG TGTGGGAAAT GGTTTCTTT AGCATTAGA CCCCATAGCC
3901 GTGCACGCCG GGGTACAGAC AGGTTCATCG ATGTTGACAT TGACTGGGAC ACCAGGTCTA
3961 TCTATTTCAT CTCCTGTCCT CTACCATAAC TCGGGACATC GGACATCTCG CTGTACCCCC
4021 CACACCCACA AAGTCTTATA AAAGCGCCAC ACCCGAGGAG GTTCGGTCGG CCCCACGAAC
4081 TCCGTGCCTC CCTGCCTGTT TACAGGGACC GAACGCTGGA GAAGCTTAGT TTCCGTGACAT
4141 CGGGCCTTACCG CGAGCAGGAA AAGGGACAGC TCATAGGCGA GGAGGGATT GAAGATGGGG
4201 ACATTTGGA TGATTGAGA GGAGGAACTA GGTACTGTAT CATGATAGTT CGGGGCAGCA
4261 TCTTGGCTGG GACATTGTTA ATACCTCGAT ATGATGAAGT GGGAGGGAGT TTTTCATGT
4321 CTTGCCAAG TCCCCACTAAT CTTTTTTTT TTTTGTACCA ACACCCAAGA TTCCGGAGAAT
4381 AGTGTAAAGGA TTGCGATTCA CAAGTGGAAAG TCTGAGGATC TTTTATATC TTTGTCTTCC
4441 GCGGACTGTT AACGATCCTA CAGCGAGCGA GCGAGCGGTC GGATGCGCTG ATCTGATAGG
4501 TGCAATATAC GGCGCTTTC TCCGGTCGTG TAGTGTAAAGC TCTGTCGGCA TAGTAGTACA
4561 CTAAAAAAAC CCTTGCATT CATGATCTGC TTGCTATTCA TTCCGAGTTA TTTCAGTGGT
4621 CACATTTCGA GATTCACAGC CATCCATCCA TATGGAAAAA TCCATTCCCA TGCTTCCCTCC
4681 CCCCCACTAT GTATGTGACC ACACGCTGCT GTCAGAATGC CAACGGTCTC AGGTACCCCTC
4741 GTCCGACTGT TTGGCATGGA GTTACATACA CTACTAGTGT AGCCCCGGGC CAAGCTACCC
4801 CGTCAAATCT ATACATATCT ATAATGGGTT TCAGGTGTT CGTTCGCTGT CAATCAAGTT
4861 TGAAACATCA CTGGGGCCGT TGGACGGTGT ATTAGACCAT TGGCTCCCTC AGCTGGCGGC
4921 TGGCGGTTG GTTCGGCAAT AACGGGACTG GACTTGAGAG GGACGAGGAG AGTCGGTTGG
4981 CTGCCTACAC TACACTACAA GCGTCCCCAC CTAACCGACG AGTCCCGTT TCCATTGTG
5041 TGCCTTAACC ATCATCTAGG GATGTCAGGG TTTGGCCGGA TCAGGGTATG TTTGGTTGAC
5101 TGTTGTCAATG TCTGATTGGG TACATATCAT GGTAGGTGTC TCGAGAACAG TAGAGTACTC
5161 GGGCCTAGCG TTTGGATGAT TACGCGAGAT ATGAGTTGTA GGCGCCCATG CAGTTGCTTG
5221 CCCATAAGCA GAAGTTGCTT TGGGATATAT TTCTCGTCTT TCAAAGGTCA CGAGGTCTG
5281 GGACGAGCGG CATGCCATC CAAAGGGTT AACATGAGAA ACCGGAATGG CCTTTGCGTT
5341 GAAATACAAA AAGTCAAGAA TAAAATCGCT TGAGGATAGG GACGTGGAAG CAAGCAAATA
5401 TGGTAAGGGA GGTACTGCTA TGTAGGTGCT CAGCAAATG CCAATTCTT GGCCCCCAAG
5461 CAGCAGTTG CTGTCAGTGC TGCTCGTGT AGCCTGGTA GTGGAACCTA AACTGCTAAC
5521 ACAGCGCAAG TGCGCATGTA AAGATATTGT GGGAGGATCT GTATGGATGG ATGAGATTAC
5581 TGCTTGGTGT TGGTTGCGAG GCACTGCCGC TGTTAGGCTT TGCTGTGCC CGTTCGACGA
5641 AGAAATACGC GGAACATATAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC
5701 GACGTAGCCG ACGGATTCTA GCAACATCCC GACTTGCTT GTAGTGTACT ATGATAGCAG
5761 CACAGTGGGG TGTTGCTCCT TGTGAGCATG GGCTCTTTT TTTTTTTCC CCCTTCCCTA
5821 GGGCGTTGAC TGGACTTGCT CTATCGTTCC CAAGGTAGGT GCCCGTCATC GATTTCCCA
5881 AGCCGTCTCC CGCCAGATTG TCGTCATAGT GTCATGATGA CCTCGGTGCG TGGGGCTGCG
5941 TGGTTACGGG GAGCTGGGAC CGCTAGGCCT CAGTGGTTGT GCCATTCAAGC GTGGGTGTGT
6001 GGAGTAGCGG TAGAGGCGCT TGGAAAGTTGT GCTAGCGGAA ACCCTGGAAT ATCTTGTACC

FIG. 7 continued

6061 CTTCGATTCC TTCTCGGGCT GCCCATGTGC TGAGGTGATG CCGGGGATCT GGCGCCAATC
6121 ATCCATTGAG GTTCCCGCAG CTTCCCGGTG CCGCGCGCG GCGCAGTGC TCACAGGACA
6181 CACCTAGACG CAGGGGCACA GGGGCACCGT TTGGTGTGCA ACTGGGTACC TGGTAGCTGT
6241 AGCAAGCACT CCACCGTCTG TGCAATCCCC CAATCCACGG CAGGAACCTA GCACCGCCGC
6301 GGCACCGAGT GAGCGAATCC ATCCGCATTG GATCCAATT CTTGCCCTG CCATCCTTCT
6361 TTCTTCCCAC TTGGCGCAAC CAACACTTCC CTTGGTCTGG GTACTCGTGT TGATCTTCAC
6421 TCTCTTTTT TCTTGGCGA CCGACTTTT ATATCCGTCC TTGCTTCCCC CTGGCCGTTG
6481 TCGTTCTTTC TACAACCTACC TTCCGTTCAT TATCCCCTT CTTGGTTGG TCGAGGACCC
6541 AAAAACAGAA CAATTCCGGC TCTTCCAGGT GGCTTGGGTG CGACTGTTA GCTCTGACC
6601 ACTAGCCGCT TACCTTCTCT TGATGTTAT ATTGGATAT CATTGAACTA CTCTTCTTG
6661 AAACGGCAGA CGAACCGAAC AGTCCCTACG GTTATTAGC GATATACGTT GTACTGATAT
6721 CCTGAGCAAG AAGAGGCAA TTATCAATTA TGCACTCTCC ATCGTCGCTG CTCATCGCAG
6781 CTCCCTTGCT CGCCAATGTA TCGGCCGAAC CGATTAGGAT ACCCCAACGC GATGTTCTCC
6841 GTGGTATCAA CATCACAGCA ACTTGCCGTT CGAGCACTAC CGAACCGGCC CAGCGGTGGA
6901 TATGCCCTG CCGTTGTAGA CTGTCCCAAG ACCAACCGA CGCTCCGGAA GGCGTGGAT
6961 TTGTCGAACG AGGAGAAGAA CTGGTTGTCG ATCCGGAGGA AGAACACCAT CCAGCCCATG
7021 AGGGACCTAC TGAAGAGGGC CAACATCACT GGGTTCGATT CCGAAACTTT CATGAATGAG
7081 GCCGCCAACCA ACGTCTCGCA ACTGCCAAT GTGCCATTG CCATTTCAGG AGGCAGGCTAT
7141 CGTGCCCTCA TGAACGGCGC CGGCTTCGTT GCTGCTGCCG ATAACCGGAT TCAAAATACC
7201 ACGGGCGCAG GTGGTATTGG AGGCTTGTG CAGTCCAGCA CATATTTGTA TGAAAACCA
7261 TGCCTTCTTG TGGTTCTTCT TATCTCGTT TCGAGTGTCA ACTGCGCCAG TTCGACGTTG
7321 GGCGGCTGTG GACGACCTG CTGGTGAACA TGTCTGGAC TCCATGCCCT TTTTTCTGTT
7381 CCCTAAAATC CCAAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAATTCGAG
7441 GACCGTGACT GTAAATTGCT AACGCAACTC TAGGGCCGGA CTTCTGGTG GTGGCTGGCT
7501 TGTCGGCACT TTGTTCTCCA ACAACTTCAG TAGCATTGAG ACCCTGCTGA GCGAGAACAA
7561 AGTCTGGGAC TTTGAGAACT CCATCTTAA AGGACCCAAG GAGGCTGGCC TTAGTACTGT
7621 CAACCGTATC CAGTACTGGT CCGAAGTGGC AAAGGAAGTT GCGAAGAAGA AGGATGCTGG
7681 CTTCGAGACA AGTATAACAG ACTACTGGGG CCGAGCATTG AGTTACCAAC TGATCGGAGC
7741 CGATATGGGC GGCCCGGCTT ACACCTTCTC CAGCATTGCC CAGACCGACA ACTTCCAGAA
7801 GGCGAAACG CCGTCCCTA TTCTGGTAGC TGACGGCCGC GCGCTGGAG ACACCATCAT
7861 CTCCCTCAAT GCTACCAACT ACGAGTTCAA CCCGTTCGAG ACGGGTAGCT GGGACCCGAC
7921 CGTCTATGGC TTTGCGCCGA CCAAGTACCT CGGCGCCAAC TTCAGCAACG GCGTGTATCCC
7981 ATCGGGAGGC AAGTGCCTG AGGGTCTCGA CCAAGCCGGC TTCGTATGG GCACCAGCAG
8041 CACGCTCTTC AACCAAGTTCC TTTGGCCAA CATCTCCAGC TACGACGGTG TTGCCAGACG
8101 TGTCATCGA GGCGGTGACT TCTGCTCTCA AGGAAATCGG CGCCAAGAGG ACGACGTCTC
8161 CCAAATCATC CCTAATCCGT TCCTGGACTG GAACAACCGG ACCAACCCCA ACGCCGACAC
8221 GCTCGAGCTC GACCTGGTCG ACGGCGCGA AGATCTGCAG AATATTCCGC TCAACCCGCT
8281 CACCCAACCC GTGCGCGCCG TCGACGTCTA CTTCGCTGTC GACTCGTCCG CCGACGTGAC
8341 AAACCTGGCCC AATGGCACCG CCCTGCGCGC CACCTACGAG CGCACTTTCG GCTCTATTT
8401 CAACGGGACA CTCTTCCCCT CGATCCCCGA CGACTGGACG TTTATAAACC TAGGCCTCAA
8461 CAACCGCCCC TCTTCTTCG GCTGCATGT TAAGAACTTT ACCTTGAACG CCAACCAAAA
8521 GGTTCCCTCC TTAATCGTCT ATGTCCCCAA CGCGCCCTAT ACCGCGCTGA GCAACGTGTC
8581 CACCTTCGAT CCGTCATACA CGATGTCTCA GCGCAACGAC ATCATCGGA ACGGATGGAA
8641 CTCAGCCACG CAGGGAAACG GCACGCTGGA TTGGAGTGG CCCACTTGC GCGCCTGCGC
8701 GGTTATCAGC AGGAGCTTAG ATCGGTTGGG CAGGCAGACG CCAGCCGCGT GCAAGACTTG
8761 CTTTGACAGG TATTGCTGGA ATGGCACAGT GAACTCCAAA GATACTGGGG TTTACATGCC
8821 TGAGTTCAAG ATTGCCTGAT CGCATGCCCT GGACTCGGGT GCTGTTGCTA TCGGAAAGAT
8881 GGTGAATGTC TGGTCGTCGG TTGTGGTGGG AGTTGTGGCG GCTACTTTGT TGTTGTAGGG
8941 GTAGGGGAGA CGTGATGATA TTCCAGTCTG ATGAAGTTGA GACTGGACTG GAGATCGCCA
9001 AGGATGCGGA GGGAAAGGAA TGCCTGGGTG TAATGTCTG ATGGATGAAG AGTCATGGAT
9061 CATGGAACGA CGGGCGGGGG ATATTGGATG ATGGATATAC CACACTGCAT GCATGCTCTA
9121 TTGATAGTAT GCTTGGCAT TTACGTTAA CAATCAATTG CTCCATCCTG ATGTTCTATC
9181 TTTTCGACA ATGGATTGAT ACTACTCCTG TTGCTCGCT CTTGAGGTTG GAAGGACTTG
9241 AGGTTGGAAG GACTTGAGGT TGTGTTCTG GAGGGAGGTT ATCGAAGTAT CATCTGTGCT

FIG. 7 continued

9301 GATGCCGATT GATAGACTGT CCTCTTCTTC GAGGCAACGA ACGGTCGGAT GAGCCTCTTT
9361 AATCATGATG CTCAGTGCCA CAAAAAGGCT CCAGCACAGC TGCCCACACC TTTCTTGCCT
9421 CGCCGTTCCCT TCCTTTTCT TTTCCCCTGT TTCCTTCTT CCTTCCATC TCATCCCGTA
9481 CCAGAGTGCC CACCGGGTAT ATATATTACC TCCTTGGCCG TTCTCCTTG ACCAATAAAT
9541 CGCTTGGTCG AGTGGCGTAA CGGTTTACCG TCTACACTTA TCACTCAAAC CAAACCAAAC
9601 CATCGAAGAA GTGACCTATC GGTCGAGGG AACGGTGATG TTCTTACGAC CAAGTTAAC
9661 CAAAGAGCGT TCCACATCGT TGAACCGTCT CCTCCAGTTG GATCTGTTA ACTTCCGCAG
9721 CGACTGAAGA AGGTATCACT TTTTTTTGG TTCCAAAAAA AAAAAAAAAA ATTAC

FIG. 8

Nucleotide sequence of the *his-3 cog^E lpl* region of linkage group I in the StLawrence wild type strain of *Neurospora crassa*. This differs from that in the Lindegren strain in many positions, summarised in figure 5. The coordinates of relevant features are given in the text. This sequence contains the weak recombinator *cog^E* and also the remnant of a transposable element *Guest* within the replaceable sequence 3' of *his-3*. StLawrence strains carry *rec-2⁺* which prevents the initiation of recombination at *cog*.

1 ACCGGGAATC GTAGCGGGCG CTAAGGCCAA GCCGCGGCAC GGGTCACTGA CCCAATGCAG
61 CGCATTCCGT CAGCAACTGA AGTGGATGTA CAAGTACATA GTAGTAGATC GCAACTGGAG
121 ATCACTCGCA CCGTGCCGCA GAACAAGGGC GACGAGCCTC AGGGCAGTTT AGCCTGCCGT
181 AACAGCACAG ACCATAGCTT ATTTCACCT GGGCGGGCGG GCGACGGCGG CACTGACATC
241 GGCAAGGCAG CATCAAGCAA CCCCTCTGTT GCTTGCCAGC TGCCGGCAA CGTCAGCGGT
301 ACAAGGAGAA ATCTGGAAGG AAAGACTTCT GGCACCGACA GGATGGCACG CGGGAAAAGT
361 TCCCAATGCA TGAGATGAGG GGCATTTGCA TTGCCTCCCG TCACCCAGTG CGAACCCCCAA
421 CCCCACCATA GCGTCTGTCG ATACATGGAG CGCGAAGTCG AGAAACCTGT AATT CCTGGT
481 AACTTTCAGG TACACAGTAC GTACTGATCC TGGTATCAA CTTGCCTGC CGAGTTTCG
541 ACGGAAAGAG GTGTGAATTG TGAAAGAGTC ATACCAAATC ACCCGATTTT CATAAAGCCC
601 GAGTCTTTTC TGTACATAAG CGACACTCGA AGCAGGGCCTC ATCTTCATAG CCTGATAGCT
661 TGTAATACTC CATCCTCGTA TCTCACTTGA CCTTGAGTTC AACCCACGT CAAACTTCAC
721 CCGACACATC GACGGATTGG GGAACAGCAC AATACCTGAA AAGCGAGAAA ACCAACACAGA
781 GGAAAACACC ATGGAGACAA CACTTCCCT CCCCTTCCTC GTCGGTGTCA GTGTTCCCTC
841 CGGACTGAAT GACATCAAGG AGGGCCTCAG CCGGGAGGAA GTCTCGTGTG TTGGCTGCGT
901 CTTCTTCGAG GTCAAGCCCC AGACCCCTGA GAAAATCCTG CGATTCTCA AGCGTCACAA
961 TGTGAATTG GAGCCCTACT TCGATGTAAC AGCCCTCGAG TCTATCGATG ATATTATCAC
1021 TCTTCTGGAC GCCGGCGCCC GCAAGGTGTT TGTCAAGACC GAGCAGTTGG CCGACCTCTC
1081 CGCATATGGC TCCCAGCTTG CCCCCATTGT CACTGGAAGC AGCGCTGCTT TGCTTCCCTC
1141 CGCCACCGAG AGCGGCCTT TGCTCTCCGG CTTCGATCAG ACTGCCTCCG AGGCTGCACA
1201 GTTCTGGAG GAGGCCAGAG ACAAGAAAAT TACCCCTTC TTCATCAAGC CCGTTCCCTGG
1261 GGCCGATCTC GAACAGTTCA TCCAGGTGCG CGCCAAGGCT AACGCCATCC CCATCCTGCC
1321 ATCCACTGGC TTGACAACAA AGAAGGACGA GGCCGGCAAG CTTGCCATCT CCACCACCTC
1381 CTCGAGCGTC TGGAAAGTCTG ACCGTCCCGA TGGTCTTCTC CCCACCGTTG TCGTTGATGA
1441 GCACGACACT GCTCTGGTC TGGTCTACAG CAGTGCCGAG AGTGTGAACG AGGCCCTCAG
1501 GACACAGACT GGTGTCTATC AGAGCCGGAA GCGCGGTCTC TGGTACAAGG GTGCTACTTC
1561 CGGAGACACT CAGGAGCTCG TCCGCATCTC GCTTGACTGC GATAACGATG CTCTCAAGTT
1621 TGTCGTGAAG CAGAAGGGTC GTTTCTGCCA CCTCGATCAG TCCGGCTGCT TTGGTCAGCT
1681 CAAAGGCCTT CCCAAGCTCG AGCAGACTTT GATTTCGAGG AAACAGTCTG CCCCCGAGGG
1741 CTCCTACACT GCCCGTCTCT TCTCCGATGA GAAGCTAGTC CGGGCCAAGA TCATGGAGGA
1801 GGCTGAGGAG CTCTGCACCG CTCAGACCCC CCAGGAAATC GCCTTGAGG CTGCCGATCT
1861 CTTCTACTTT GCTCTTACCA GGGCCGTTGC TGCCGGCGTT ACTCTTGCCG ATATCGAAAG
1921 GAGCCTTGAC GCCAAGAGCT GGAAGGTCAA GCGCAGGACT GGAGATGCTA AGGGTAAGTG
1981 GGCTGAGAAG GAGGGCATCA AGCCTGCGGC GTCCGCTCTC GCTGCCACTT CGGCCCTGT
2041 CACCAAGGAG GCCGCCAGG AGACCACCCC TGAGAAGATC ACCATGAGAC GTTTCGACGC
2101 CTCCAAGGTC TCTACCGAGG AGCTCGATGC TGCTCTCAAG CGTCCTGCGC AAAAGTCGTC
2161 CGATGCCATC TACAAGATCA TTGTCCCCAT CATCGAGGAC GTCCGCAAGA ACGGCGACAA
2221 GGCTGTTCTG TCGTACACTC ACAAGTTCGA GAAGGCTACC TCTCTTACTA GCCCCGTCCT
2281 GAAGGCGCCC TTCCCCAAGG AGCTTATGCA GCTCCCTGAG GAGACCATTG CTGCCATCGA
2341 CGTGTCTTC GAGAACATCC GCAAGTTCCA CGCCGCCAG AAGGAGGAGA AGCCCCTCCA
2401 GGTGAGGACG ATGCCCGGTG TTGTCTGCAG CCGTTTCTCT CGTCCCATCG AGGCCGTGG
2461 CTGCTACATC CCCGGCGGTG CCGCCGTTCT CCCCAGCACT GCCCTTATGC TGGGTGTTCC
2521 CGCCATGGTC GCCGGCTGCA ACAAGATTGT GTTGCCTCT CCTCCCCGCG CCGACGGAAC
2581 CATCACTCCC GAGATTGTCC ACGTCGCTCA CAAGGTTGGG GCCGAGTCCA TCGTGCTTGC
2641 CGGGGGTGCC CAGGCCGTAG CTGCCATGGC CTACGGCACC GAGAGCATCA CCAAGGTCGA
2701 CAAGATTCTC GGCCCCGGTA ACCAGTTCGT CACTGCTGCC AAGATGTTCG TCAGCAACGA

FIG. 8 continued

2761 CACCAACGCT GCCGTTGGTA TTGACATGCC CGCTGGCCCG TCCGAGGTGC TGGTCATCGC
2821 TGACAAGGAC GCCAACCCCG CGTCGTTGC CTCGGATCTC CTGTCCCAGG CTGAGCACGG
2881 CGTTGACAGT CAGGTACATCC TGATCGCTAT TGACCTCGAC GAGGAGCATC TTCAGGCTAT
2941 TGAGGACGAG GTTCACCGTC AGGCTACGGA GCTTCCTCGC GTCCAGATTG TCCGTGGCTC
3001 CATCGCCCAC TCGATCACCG TGCAAGTCAA GACCCTCGAG GAGGCCATGG AGCTCAGCAA
3061 CAAGTACGCT CCTGAGCACT TGATCCTCCA GATCAAGGAG GCCGAGAAGG CTGTCGATCT
3121 TGTCACTAAC GCCGGTAGTG TCTTCATTGG CGCCTGGACT CCTGAGTCCG TTGGCGATTA
3181 CTCTGCTGGT GTTAACCACT CGCTGCGTAA GTTACATATC ATAAATAGCC CCGCTTCACA
3241 GATTCTTCTG CTAACGTCAA GACACATAGC TACCTATGGC TTTGGCAAGC AGTACTCTGG
3301 CGTCAATTTC GCCTCGTTCG TCAAGCACAT TACCAAGCTCC AACTTGACTG CCGAGGGTCT
3361 CAAAAACGTC GGCCAGGCTG TCATGCAGTT GGCTAAGGTT GAGGAGCTCG AGGCTCACAG
3421 AAGGGCGGTC AGCATCCGTC TTGAGCACAT GAGCAAGAGC AACTAAACGG AAATTCTTT
3481 CGAAGTAGCA AAAAAAAA AAAAAACAA GAACAAAAGG ATGTAGTGGG TTGATGTATA
3541 TCTGGGTCA TTTGGGCACA TAGAGTAATG ATAACGAGTT TTGGACATTG TACTGTTCTG
3601 TACAGGCTGA AGATCAGTAC ATGAATCTGT TGGTAAGTGT GGAGACCCAA ACgtcccttg
3661 AGTTTTCTC CCTATTCCAG AGGTGCTCGT CCCTGGGTGT TTATTTCAT TATTACATCA
3721 ACCTTTTTT TTTTTTTTT TTTTCAGAT CATGCGTACA TGAACGGGGG AAGCACAGAC
3781 GATCGAAACG TGGATGTCAC AATGTCGCTG CAGTGATGCT GCATGCGAT AAGCGCCCAT
3841 CTCAATATAC TTGCACTCTT GCACGTTGCA TGTGAACCTC CCAAACAACC GAATAAAAGA
3901 CGCGAAAAAA TGAAGATAAA AAAAAACCAT AAAAAAAATC AGAGGGAGTG TGGGAAATGG
3961 TGTCTTTAG CATTAGACCC CCATAGCCGT GCACGCCCG GTACAGACAG GTTCATCGAT
4021 GTTGACATTG ACTGGGACAC CAGGTCTATC TATTTTATCT CCTGTCCTCT ACCATACATC
4081 GGGACATCGG ACATCTTGCT GTACCCCCCA CACCCACAAA GCCTTATAAA AGCGCCACAC
4141 CCGAGGAGGT TCGGTCGGCC CCACGAACTC TGTGCCTCCC TGCCTGTTA CAGGGACCGA
4201 ACGCTGGAGA ATCTTACTAG TTTCCTGACA TCCGGCCTAC CCGAGCAGGA AAAGGGACAG
4261 CTCATAGGCG AGGAGGGATT TGAAGATGGG AACATTTGG GTGATTGAG AGGAGGAAC
4321 AGGTACTGCA TCATGATAGT TCGGGGCAGC ATCTGGCTG GGACATTGTT AATACCTCGA
4381 TATGATGAAG TAGGAGGGAG TTTTGCCTG TCTTGCCTGAA GTCCAGAGAT CTGTTTATT
4441 TTATTTTTA TGGATGTAGT GTATCAACAC CCAAGATTG GAGAATAGTA CTAGGATTG
4501 CATTACAAG TGGAAAGTCTT GAGAATCGTT GTATATCCTT GTCTCCTCG GAATGTTAAC
4561 AATCCTACAG CGAGCGAGCG AGCGCTCGGA TGCGCTGATC TGATAGGCGC AATATACGGC
4621 CGCTTCTCC GGTCTGTAG TGTAAGCTCT GTGGGCATAG TACACTAAA AAACCTTGC
4681 ATTTCATGAT CTGCCTGCTA TTCATTCCGA GCTATTCAG TGGTCACATT TCGAGGAAGA
4741 AAGAAAGCAA CTAAGATTCA CAGCCATCCA TCCATCCATA TGGAAGAATA ATCCATTCCC
4801 ATGTTCCCTC CCCCCCACTA TGTATGTGAC CACACGCTGC TGTCAGAATG CCAACGGTCT
4861 CAGGTACCCCT CGTCCGACTG TTTGGCATGG AGTTACATAC ACTACTAGTG TAGCCCCGGG
4921 CCAAGCTACC CCGTCAAATC TATACATATC TATAACGGGT TTCAGGGTT TCGTTCGCTG
4981 TCAATCAAGT TTGAAACATC ACTGGGGCCG TTGGACGGTG TATTAGACCA TTGGCTCCCT
5041 CAGCTGTTG GCGGCTGGGC GGCTGGGTCA AACGGCAATA ACGGGACTCG AGAGGGACGA
5101 GGAGAGTCGG TTGGCTGGCT GCAATACAAG CGTTCCCACC TAACCAACGA GTCCCGTTT
5161 CCATTTGTGT GCCTAACCAT CATCTAGGGA TGTCAGGGTT TGGCCGGATC AGGGTATGTT
5221 TGGTTGACTG TTGTCATGTC TGATTGGTA CATATTATGG TAGGTGTCTC GAGAACAGTA
5281 GAGTACTCGG GCCTAGCGTT TGGATGATTA CGCGAGATAT GAGTTGTGGG CCGCCATGCA
5341 GTTGCTTGTC CATAAGCAGA AGTTGCTTG GGATATATT CTCGTCTTTC AAAGGTCACG
5401 AGGTCCCTGGG ACGAACGGCA TCGCCATCCA AAGGGTTGAA CATGAGAAAC CTGAATGGCC
5461 TTTGCCTGTA AATACAAAAA GTCAAGAACAA AAATCGCTTG AGGATAGGGG CGTGGAAAGCA
5521 AGCAAATATG GTAAGAGAGG TATACATCAA CCCTGGTTCA ATTGTTAGCG TGGTTCTTCC
5581 TCCACGTCT CGTTCATGAC GGTAAACAGT ACCAGGCTAA CAATTAAACC AGGGTTGATG
5641 TGTACTGATA TGTAGGTGCT CAGCAAATG CCAATTCTT TGGCCCCAAG CAGCAGTTG
5701 CTGTCAGTGC TGCTCGTGTGTC AGCCTGGTA GTGGAACCTA AACTGCTAAC ACAGCGCAAG
5761 TGCGCATGTA AAGATATTGT GGGAGGGATCT GTATGGATGG ATGAGATTAC TGCTTGGTGT
5821 TGGTTGCGAG GCACTGCAGGC TGTTAGGCTT TGCTGTGCC CGTTCGACGA AGAAATACGC
5881 GGAACATAAA ATTGGATACC TAGACTTACT GCCTATGGGA GGTATCTACC GACGTAGCCG
5941 ACGGATTCTA GCAACATCCC GACTTGCTT GTAGTGTACT ATGATAGCAG CACAGTGTG

FIG. 8 continued

6001 CTCCTTGTGA GAATGGGCTC TTTTTTTTT TCCCCCTTCC CTAGGGCGTT GACTGGACTT
6061 GCTCTATTGT TCCCAAGGTA GGTGCCCGTC ATCGATTTTC CCAAGTCTCC CGCCAGATTG
6121 TCGTCATAGT GTCATGATGA CCTCGGTGCG TGGGGCTGCG TGTTACGGG GAGCTGGGAC
6181 CGCTAGGCCT CAGTGGTTGT GCCATTCAAC GTGGGTGTGT GGAGTAGCGG TAGAGGCCT
6241 TGGAAAGTTGT GCTAGCGGAA ACCCTGGAAT ATCTTCTACC CTCGATTCTT TCTCGGGCTG
6301 CCCATGTGCT GAGGTGATGC CGGGGATCTG GCGCCAATCA TCCATTGAGG TTCCCGCAGC
6361 TTCCCGGTGC CGCGCGCGGG CGCAGTTGCT CACAGGACAC ACCTAGACGC AGGGCACAG
6421 GGGCACCGTT TGGTGTGCAA CTGGGTACCT AGCTGTAGCA AGCACTCCAC CGTCTGTGCA
6481 ATCCCCAAT CCACGGCAGG AACTTCGCAC CGCCGCGGCA CCGAGTGAGC GAATCCATCC
6541 GCATTGGATC CCAATTCTTG CCCTTGCCAT CCTTCTTCT TCCCACTTGG CGCAACCAAAC
6601 ACTTCCCTTG GTCTGGGTAC TCGTGTGAT CTTCACTCTC TTTTTTCTT GGGCGACCGA
6661 CTTTTATAT CCGTCCTTGC TTCCCCCTGG CCGTTGTCGT TCTTCTACA ACTACCTTCC
6721 GTTCATTATC CCCTTCTTG GTTCGGTCGA GGACCCAAAA ACAGAACAAAT TCCGGCTCTT
6781 CCAGGTGGCT TGGGTGCGAC TGTTTAGCTC TTGACCACTA GCCGCTTACC TTCTCTTGAT
6841 GTTTTATTT GGATATCATT AAACACTCT TTCTTGAAAC GGAGACGAA CGGAACAGTT
6901 CCTACGGTAT ATTAGCGATA TACGTTGTAC TGATATTCTG AGCAAGAAGA GGCAAATTAT
6961 CAATTATGCA TCTCCCTTCG TCGCTGCTCA TCGCAGCTCC CTTGCTCGCC AATGTATCGG
7021 CCGAACCCAT TAGGATACCC CAACGCGATG TTCTCCGTGG TATCAACATC ACAGCAACTT
7081 GCCGTTCGAG CACTACCGGA TTCGCCAGC GGTGGATATG CCCCTGCCGT TGTAGACTGT
7141 CCCAAGACCA AGCCGACGCT CGGAAAGGCC GTGGATTGT CGAACGAGGA GAAGAACTGG
7201 TTGTCGATCC GGAGGAAGAA CACCATCCAG CCCATGAGGG ACCTCCTGAA GAGGGCCAAC
7261 ATCACTGGGT TCGATTCCGA GACATTATG AATGAGGCCG CCAACAACAT CTCGCAACTG
7321 CCCAATGTCG CCATTGCCAT TTCAGGAGGC GGCTATCGT CCCTCATGAA CGGCGCCGGC
7381 TTCGTTGCTG CTGCGGATAA CCGAATTCAA AATACCACGG GCGCAGGTGG TATTGGAGGC
7441 TTGTTGCACT CCAGCACATA TTTGTATGTA AAGTGGTTCT TCTTATCTCG TTTTCGAGTG
7501 TCAACTGCGC CAGTTCAGAG TTGGCGGGCT GTGGACGACC TTGCTGGTGA ACATGTCTTG
7561 GACTCCATGC CCCTCTTCG TTTCTCAA TCAAGAAGTC GAGGACCGTG ACCGTAATAC
7621 GCTAACGCAA CTCTAGGGCC GGACTTCTG GTGGTGGCTG GCTTGTGCGC AGTTGTTCT
7681 CCAACAACTT CAGCAGCATT GAGACCCCTGC TGAGCGAGAA CAAAGTCTGG GACTTGAGA
7741 ACTCCATCTT TAAAGGGCCC AAGGAGGCTG GCCTAGTAC TGTCAACCGC ATTCACTACT
7801 GGTCCGAAGT GGCAAAGGAA GTGCCAAGA AGAAGGATGC TGGCTTCGAG ACAAGTATAA
7861 CAGACTACTG GGGCCGAGCA TTGAGTTACC AACTGATCGG AGCCGATATG GGCAGGCCGG
7921 CTTACACCTT CTCCAGCATT GCCCAGACCG ACAACTTCCA GAAGGCCGAA ACGCCGTTCC
7981 CTATTCTGGT AGCTGACGGC CGCGCGCCTG GAGACACCAT CATCTCCCTC AATGCTACCA
8041 ACTACGAGTT CAACCCGTT GAGACGGTA GCTGGGACCC GACCGTCTAT GGCTTGCAGC
8101 CGACCAAGTA CCTCGGCGCC AACTCAGCA ACGGCGTGAT CCCATCGGA GGCAAGTGCAG
8161 TTGAGGGTCT CGACCAAGCC GGCTTCGTCA TGGGCACCAG CAGCACGCTC TTCAACCAAGT
8221 TCCTTTGGC CAACATCTCC AGCTACGACG GTGTTGCCCG ACGTGTCTAT CGAAGCCGTG
8281 ACTTCTGTCC TCAAGGAAAT CGGCGCCAAG AGGACGACGT CTCCCAAATC ATCCCTAAATC
8341 CGTTCTGGA CTGGAACAAAC CGGACCAACC CCAACGCCGA CACGCTCGAG CTCGACCTGG
8401 TCGACGGCGG CGAAGATCTG CAGAATATTG CGCTCAACCC GCTCACCAA CCCGTGCGCG
8461 CCGTGGACGT CATTTCGCT GTCGACTCGT CCGCCGACGT GACAAACTGG CCCAATGGCA
8521 CGGCCCTGCG AGCCACCTAC GAGCGCACTT TCGGCTCTAT TTCCAACGGG ACACTCTTCC
8581 CCTCGATCCC CGACGACTGG ACGTTATAA ACCTAGGCCT CAACAACCGC CCCTCTTCT
8641 TCGGCTGCGA TGTAAAGAAC TTTACCTTGA ACGCCAACCA AAAGGTTCCC CCCTTAATCG
8701 TCTATGTCCC CAACCGGCC TATAACCGCG TGAGCAACGT GTCCACCTTC GATCCGTCT
8761 ACACCATGTC TCAGCGAAC GACATCATCG GCAACGGATG GAACTCAGCC ACGCAGGGAA
8821 ACGGCACGCT GGATTCGGAG TGGCCCACTT GCGTCGCCTG CGCGGTTATC AGCAGGAGCT
8881 TAGATCGGTT GGGCAGGCAG ACGCCAGCCG CGTGAAGAC TTGCTTGAG AGGTATTGCT
8941 GGAATGGCAC AGTGAACCTCA AAAGATACAG GGGTTACAT GCCTGAGTTC AAGATTGCGG
9001 ATGCGCATGC CCTGGACTCG GGTGCTGTTG CTATCGAAA GATGGTGAAT GTCTGGTCGT
9061 CGGTTGTGGT GGGAGTTGTG GC GGCTACTT TGTTGTTGTA GGGTAGGGG AGACGTGATG
9121 ATATTCCAGT CTGATGAAGT TGAGACTGGA CTGGAGATCG CCAAGGATGC GGAGGGAAAG
9181 GAATGCGTGG TGTAAATGTC ATGATGGATG AAGGGTCATG GATCATGGAA CGACGGGGCG

FIG. 8 continued

9241 GGGATATTGG ATGATGGATA TACCACACTG CATGCATGCT CTATTGATAA TATGCTTG
9301 CATTACGTT TAACAATCAA TTGCTCCATC CTGATGTTCT ATCTTCGAC ACTGGATTGA
9361 TACTACTCCT GTTGCTTCCC TCTTGAAGTT GGAAGGACTT GAGGTTGGAA GGACTTGAGG
9421 TTGTTTGTTC TGAGGGAGGT TATCGAAGTA TCATCTGTGC TGATGCCGAT CGATAGACTG
9481 CCCTCTTCTT CGAGGCAACG AACGGTCGGA TGAGCCTCTA ATCATGATGC TCAGTGCCAC
9541 AAAAAGGCTC CAGCACAGCT GCCCACACCT TTTTGCCCTC GTCGCTCCTT CCTTTTTTC
9601 CCCCCCTTTC TTCCTTCCA TCTCATCCCG TACCAGAGTG CCCACCGGGT ATATATATTA
9661 CCTCCTTGGC CGTTCTCCTT TGACCAATAA ATCGCTTGGT CGAGTGGCGT AACCGTTTAC
9721 CGTCTACACT TATCACTCAA ACCAAACCAA ACCATCGAAG AAGTTACCTA TCGGTTCGAG
9781 GGAACGGTGA TGTTCTTACG TTCAAGTTAA CCCAAAGAGC GTTCCACATC GTTGAACCGT
9841 CTCCTCCAGT TCTTGGATCT GTTTAACTTC CGCAGCGACT GAAGAAGTAA TCACTTTTT
9901 TTTTTTGGT TCCAAAAAAA AAAAAAAA TTAC

FIG. 9

Construction of the components of the sequence diversification cross: Parent (variant 1) and Parent (variant 2). For convenience, plasmid sequences are shown as linear. The cross hatched region in the chromosome is dispensable. Stippled sequence in the plasmid indicate the multiple cloning site for inserting foreign DNA. Crossovers in region 1 and region 2 insert the foreign sequence to be diversified into chromosome 1 of *Neurospora crassa* adjacent to the recombination hotspot *cog^L*. Parent (variant 2) containing a version of the foreign sequence with multiple differences from that in parent (variant 1) is similarly constructed. Parent (variant 1) and parent (variant 2) are crossed and conversion events (stippled arrow) initiated (X) at *cog^L* recombine the sequence differences in variant 1 and variant 2 to form new combinations. Sequences are identical except for those that distinguish variant 1 and variant 2. *rec-2* on linkage group V permits *cog^L* to be active. For simplicity, genes not directly related to the diversification are omitted. See text for further details.

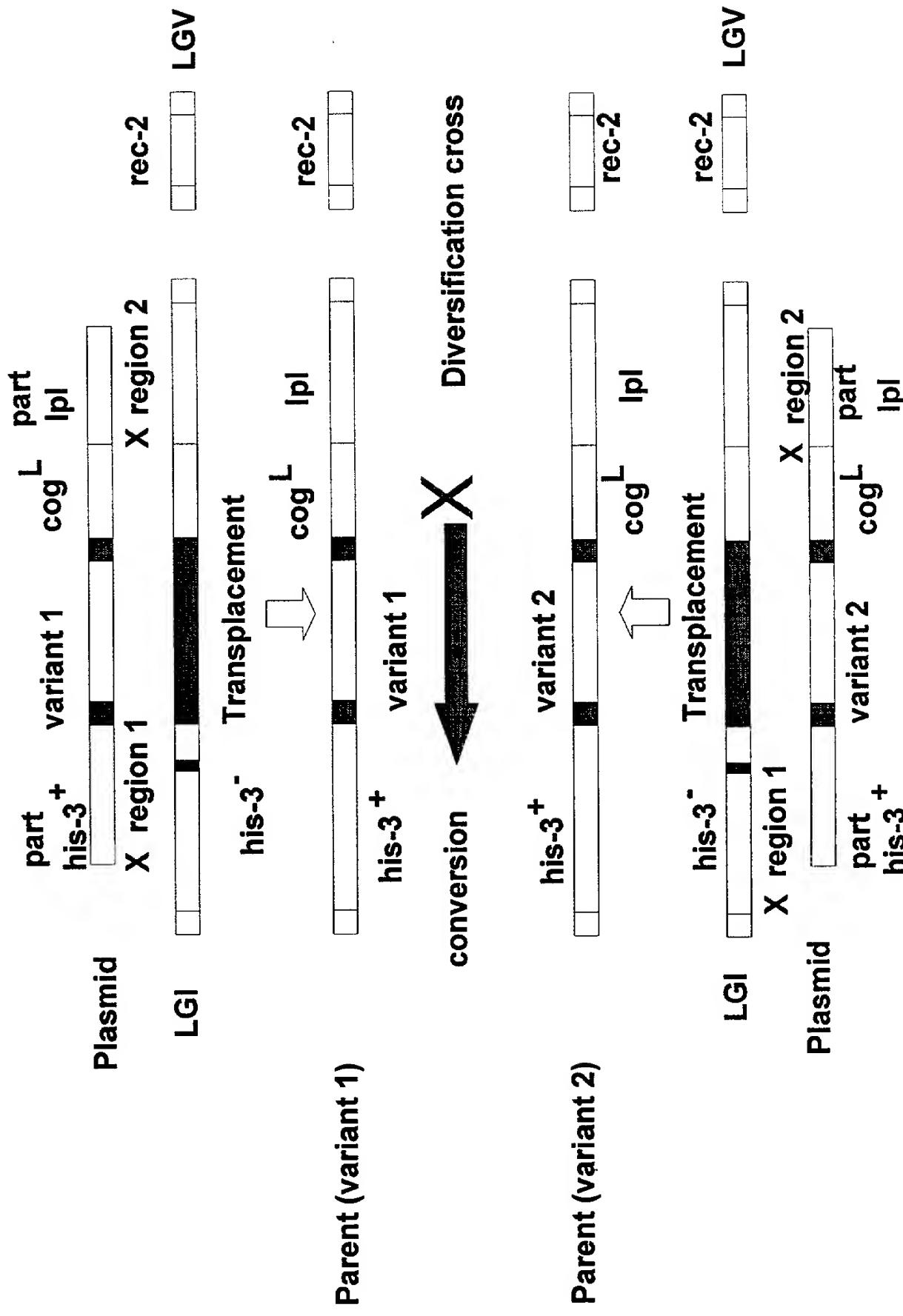


FIG. 10

Construction of parent (variant 1) and parent (variant 2) enabling selection of progeny that have experienced conversion in the foreign DNA. Complementing pairs of *his-3* alleles are used to obtain *his-3* alleles that have experienced conversion in the foreign DNA. Parent (variant 1) and parent (variant 2) are crossed and *his⁺* recombinants are selected. These must all have experienced conversion events affecting the foreign DNA since the events begin at *cog^L*. The *his-3* alleles in parent (variant 1) and parent (variant 2) are non complementing to ensure that selection yields recombinants and not aneuploid progeny having two copies of all or part of linkage group I.

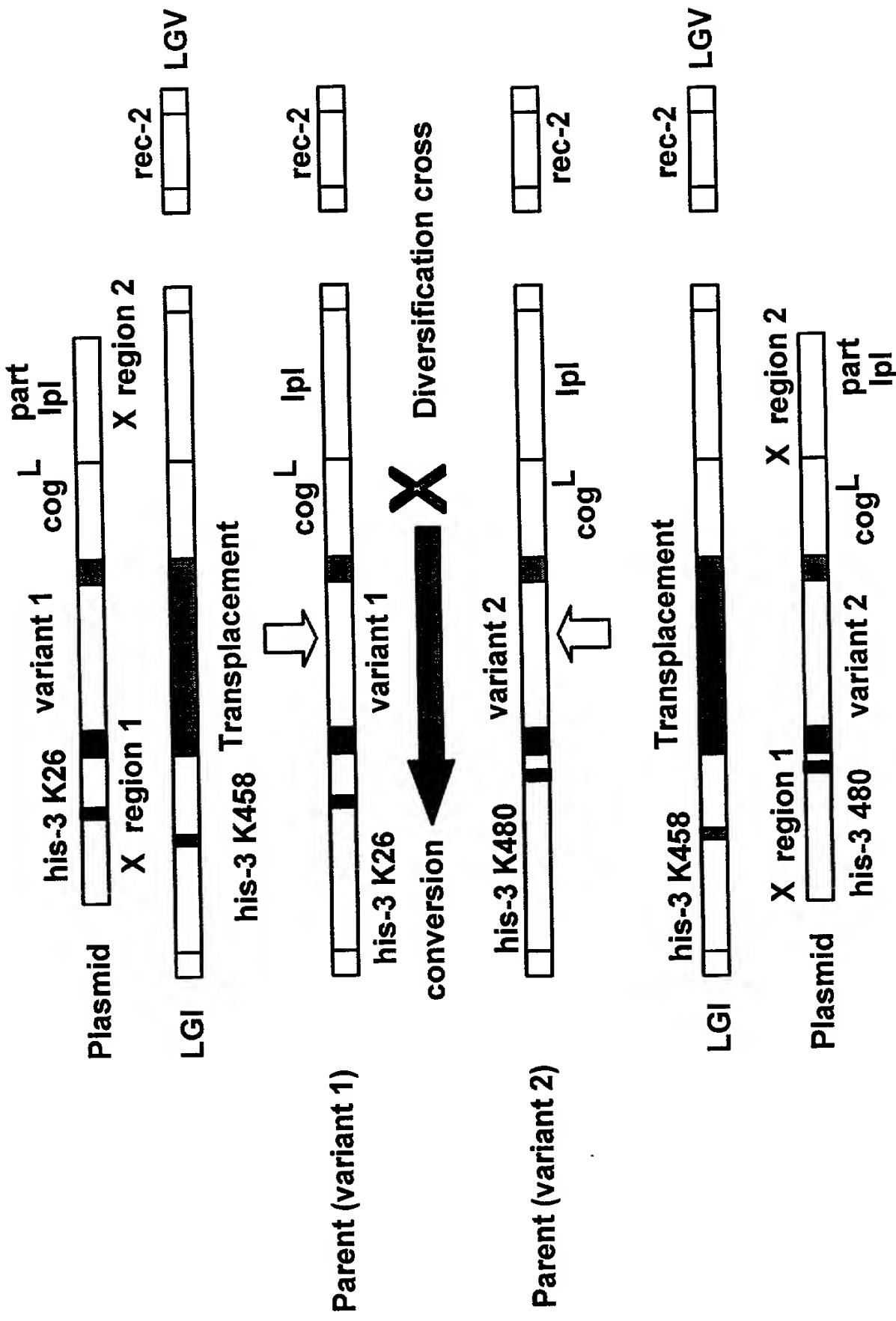


FIG. 11

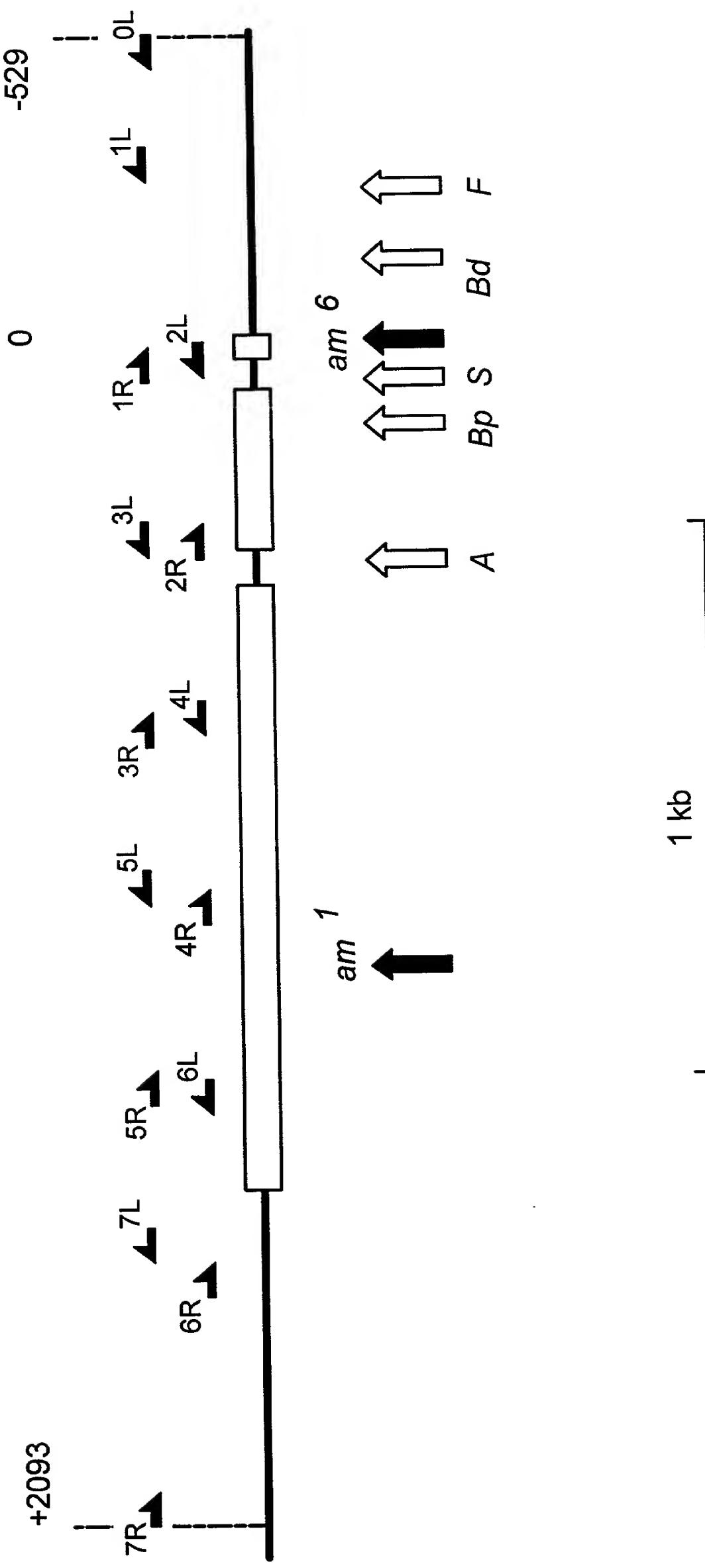


FIG. 12

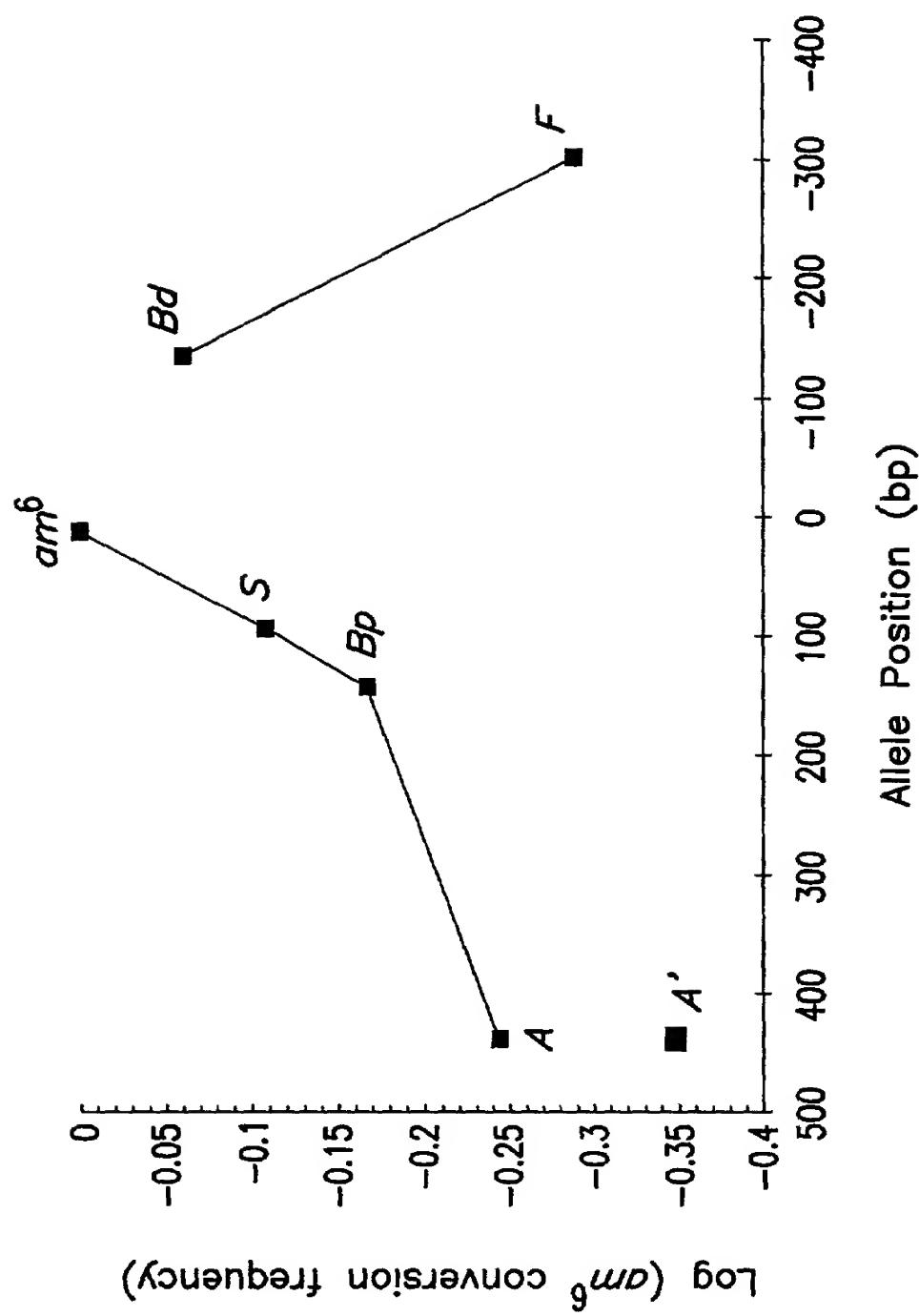


FIG. 13

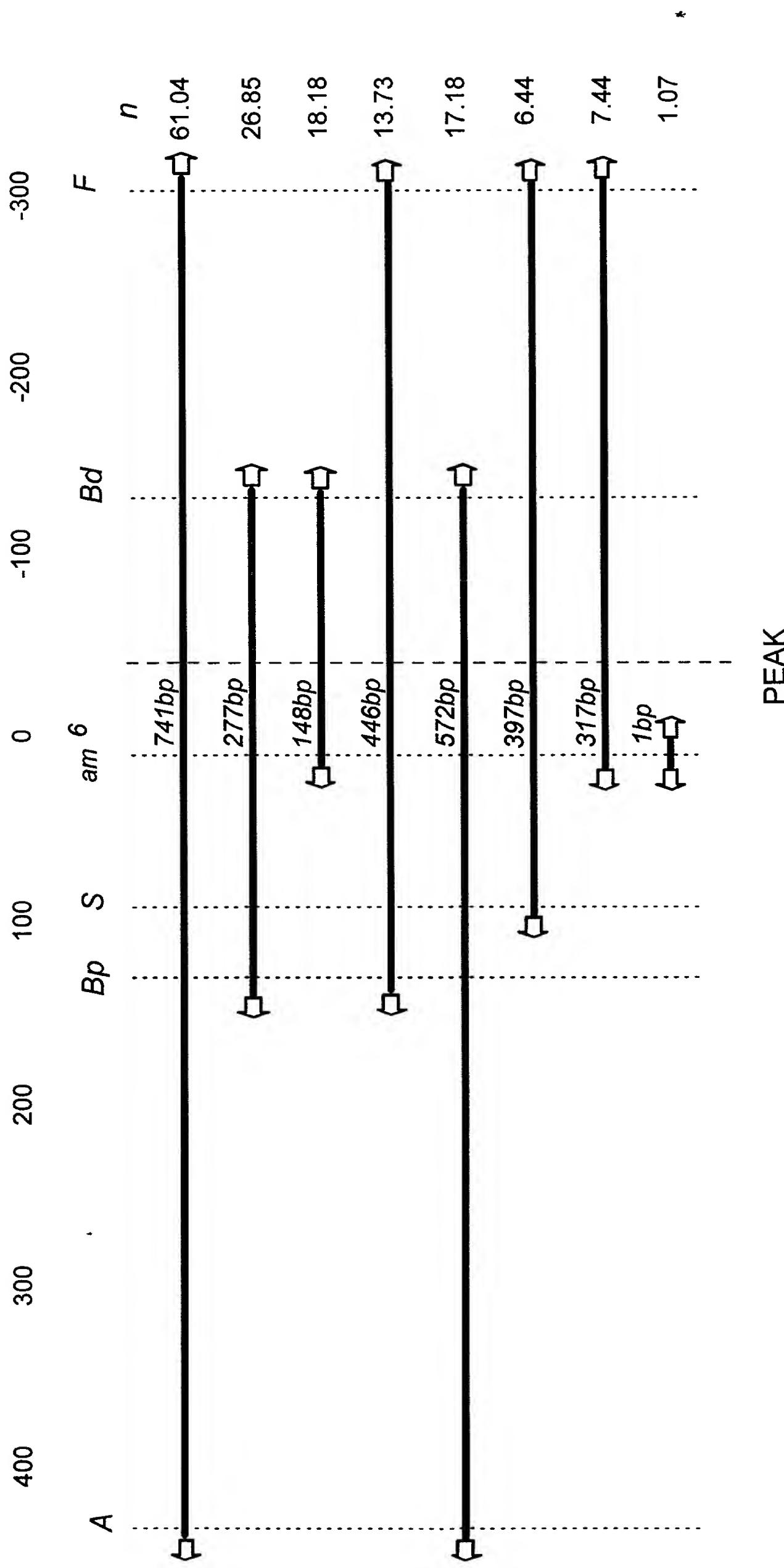


FIG. 14

